

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2003, 07:27:37 ; Search time 21900 Seconds
(without alignments)
11635.893 Million cell updates/sec

Title: US-09-913-159a-10

Perfect score: 6229

Sequence: 1 ctgacgcgcgcctagtcgcgc.....attcccccgaagaatgcac 6229

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_pro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6229	100.0	6229	6 AX032758	AX032758 Sequence
2	2676.2	43.0	9709	6 AR067813	AR067813 Sequence
3	2676.2	43.0	9709	6 AR224437	AR224437 Sequence
4	2676.2	43.0	9709	6 AX032749	AX032749 Sequence
5	2676.2	43.0	9709	6 AX032749	AX032749 Sequence
6	2676.2	43.0	14824	12 AF324493	AF324493 HIV-1 vec
7	2665	42.8	9000	14 HIV1UD6942	U26942 HIV-1 vec
8	2610.6	41.9	9754	14 HIV2132	D86069 Human immun
9	2607.4	41.9	9699	14 AF070521	AF070521 HIV-1 E9
10	2599.4	41.7	9213	6 E00987	E00987 Genomic DNA
11	2599.4	41.7	9213	6 I04549	I04549 Sequence 11
12	2599.4	41.7	9752	14 HIVMCK1	D86068 Human immun
13	2589.4	41.7	9770	14 HIVPV22	K02083 Human immun
14	2598.2	41.7	3156	6 E01088	E01088 Nucleic ac1
15	2597.8	41.7	8932	14 HIVBH102	M15654 Human immun
16	2597.8	41.7	8933	6 AR094659	AR094659 Sequence
17	2597.8	41.7	8933	6 AX078307	AX078307 Sequence
18	2597.8	41.7	8933	6 AX078308	AX078308 Sequence
19	2597.8	41.7	8933	6 AX078313	AX078313 Sequence
20	2597.8	41.7	9748	14 REHTLV3	X01762 Human T-cell
21	2597.8	41.7	9749	6 I07983	I07983 Sequence 1
22	2596.2	41.7	9748	6 E01099	E01099 DNA sequence
23	2595	41.7	3156	14 HIVHXB3	M14100 Human immun
24	2593	41.6	9181	6 AX074066	AX074066 Sequence
25	2593	41.6	9181	14 AF033819	AF033819 HIV-1, CO
26	2593	41.6	9719	6 AR268768	AR268768 Sequence
27	2593	41.6	9719	14 HIVHXB2CG	K03455 Human immun
28	2583.4	41.5	9795	14 HIVTH475A	L31963 Human immun
29	2574.4	41.3	9193	6 A04321	A04321 Human immun
30	2574.4	41.3	9229	14 HIVBRUCG	K02013 Human immun
31	2572.4	41.3	9193	6 A07867	A07867 Human immun
32	2572.2	41.3	9609	14 HIVYU12055	U12055 Human immun
33	2563.4	41.2	2565	6 BD000747	BD000747 Recombembl
34	2560.4	41.1	2565	6 AX166280	AX166280 Sequence
35	2558.4	41.1	12479	6 AR287304	AR287304 Sequence
36	2558.4	41.1	12494	6 AR140325	AR140325 Sequence
37	2558.4	41.1	12494	6 BD105983	BD105983 Animal mo
38	2554.6	41.0	3563	6 AR094661	AR094661 Sequence
39	2554.6	41.0	3563	14 HIVH3BH8	K02011 HIV-1 Isola
40	2553	41.0	9781	14 HIVF12CG	Z11530 Human immun
41	2543.6	40.8	2918	6 I05801	I05801 Sequence 2
42	2543.4	40.8	9091	6 A00647	A00647 Lymphadenop
43	2543.4	40.8	9091	6 BD165833	BD165833 DNA obtai
44	2543.4	40.8	9094	6 A10447	A10447 Complete sy
45	2518.6	40.4	2565	14 HIVGPI60BN	L42371 Human immun

ALIGNMENTS

RESULT 1
LOCUS AX032758 6229 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 10 from Patent WO0047223.
ACCESSION AX032758
VERSION AX032758.1 GI:10279735
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 Schreiber, M.
AUTHORS Viral vaccine
TITLE Patent: WO 0047223-A 10 17-AUG-2000;
JOURNAL SCHREIBER MICHAEL (DE); STRATHMANN AG & CO (DE)

```
FEATURES
    source          Location/Qualifiers
                    1..6229
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
                    /note="Synthetische"
                    sig_peptide
                    /note="env ATG"
                    /note="env ATG"
                    misc_feature
                    /note="env AGT, gp120 Anfang"
                    /note="BstEII-Schnittstelle"
                    misc_feature
                    /note="BamHI-Schnittstelle"
                    /note="env TAA, Stop"
BASE COUNT      1795 a 1332 c 1518 g 1584 t
ORIGIN
Query Match      100.0%; Score 6229; DB 6; Length 6229;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 CTGACGGCCCTGTAGCGGCGCATTAAAGCGGCGGTGTGTTACGCGCAGCGTGA 60
DB      1 CTGACGGCGCCTGTAGCGGCGCATTAAAGCGGCGGTGTGTTACGCGCAGCGTGA 60
QY      61 CCGGTACACTGTGCAGGCGCCCTAGCGCCGCTTCCTTCGCTTTCCTCCCTCTTTCG 120
DB      61 CCGGTACACTGTGCAGGCGCCCTAGCGCCGCTTCCTTCGCTTTCCTCCCTCTTTCG 120
QY      121 CCAGCTTGCGCGCTTTCGCCGTCAAGCTTAAATCGGGGCGCTCCCTTAGGGTTCGAT 180
DB      121 CCAGCTTGCGCGCTTTCGCCGTCAAGCTTAAATCGGGGCGCTCCCTTAGGGTTCGAT 180
QY      181 TTATGCTTTAGCGCACCCTGACCCCAAAAACCTTGATAGGGTGATGTTACAGTAGTG 240
DB      181 TTATGCTTTAGCGCACCCTGACCCCAAAAACCTTGATAGGGTGATGTTACAGTAGTG 240
QY      241 GGCCATGCGCCTGTATAGAGGTTTTCGCCCTTTCGAGCTTGAGATCCAGTCTTTTATA 300
DB      241 GGCCATGCGCCTGTATAGAGGTTTTCGCCCTTTCGAGCTTGAGATCCAGTCTTTTATA 300
QY      301 GTGACCTCTGTGTTCCAACTGGAACAACACTCAACCTATCTCGGTATATCTTTGAT 360
DB      301 GTGACCTCTGTGTTCCAACTGGAACAACACTCAACCTATCTCGGTATATCTTTGAT 360
QY      361 TATAGGGAATTTGCCATTTCCGCTATTCGTTAAAAATGAGCTGATTTTAAACAAAT 420
DB      361 TATAGGGAATTTGCCATTTCCGCTATTCGTTAAAAATGAGCTGATTTTAAACAAAT 420
QY      421 TTAACGGGAATTTTAAACAAATATTAACGCTTACATTTCCATTGCGCATTCAGGCTGCG 480
DB      421 TTAACGGGAATTTTAAACAAATATTAACGCTTACATTTCCATTGCGCATTCAGGCTGCG 480
QY      481 CAACCTGTTGGGAAGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGGGAAAG 540
DB      481 CAACCTGTTGGGAAGGCGATCGGTGCGGCGCTCTTCGCTATTACGCCAGCTGGGAAAG 540
QY      541 GGGATGTGCTGCAAGGGGATTAAGTTGGGTAAAGCCAGGGTTTCCAGTACAGACGTG 600
DB      541 GGGATGTGCTGCAAGGGGATTAAGTTGGGTAAAGCCAGGGTTTCCAGTACAGACGTG 600
QY      601 TAAACGACGCGCAGTAGAGCTAGTATTAATAGTAATCAATTAACGGGGTCAATTAAT 660
DB      601 TAAACGACGCGCAGTAGAGCTAGTATTAATAGTAATCAATTAACGGGGTCAATTAAT 660
QY      661 CATAGCCCATATATGAGATTCGCGGTACATTAACCTTAAGTAAATGGCGCCGCGGTGA 720
DB      661 CATAGCCCATATATGAGATTCGCGGTACATTAACCTTAAGTAAATGGCGCCGCGGTGA 720
QY      721 CCGGCCAAGACGCCCGCCCATTTAGCTCAATTAATGACGTATGTTCCCATAGTAACGCCA 780
DB      721 CCGGCCAAGACGCCCGCCCATTTAGCTCAATTAATGACGTATGTTCCCATAGTAACGCCA 780
```

```
DB      721 CCGGCCAAGACGCCCGCCCATTTAGCTCAATTAATGACGTATGTTCCCATAGTAACGCCA 780
QY      781 ATAGGACTTTCCATTGACGCTCAATGGGTGAGATTTACGGTAAACTGCCACTTGGCA 840
DB      781 ATAGGACTTTCCATTGACGCTCAATGGGTGAGATTTACGGTAAACTGCCACTTGGCA 840
QY      841 GTACATCAAGGTATCATATATGCGCAAGTAGCGCCCTTATGCGCAATGACGTAATGG 900
DB      841 GTACATCAAGGTATCATATATGCGCAAGTAGCGCCCTTATGCGCAATGACGTAATGG 900
QY      901 CCCGCTGGCATTAATGCCCATACATGACCTTATGGAGCTTTCCTACTTGGCAGTATAC 960
DB      901 CCCGCTGGCATTAATGCCCATACATGACCTTATGGAGCTTTCCTACTTGGCAGTATAC 960
QY      961 TACGATTTAGTATGCTATTTACATGCTATGATGCGGTTTGGAGTATCATATGGCGCT 1020
DB      961 TACGATTTAGTATGCTATTTACATGCTATGATGCGGTTTGGAGTATCATATGGCGCT 1020
QY      1021 GGATAGCGGTTTGCATCAGGCGGATTTCCAAAGTCTCCACCCCATTTGACGTCATGGAGT 1080
DB      1021 GGATAGCGGTTTGCATCAGGCGGATTTCCAAAGTCTCCACCCCATTTGACGTCATGGAGT 1080
QY      1081 TTGTTTGGCACAACAAATCAACGGGACTTTCACAAATGTGTAACAACCTCCGCCATTTG 1140
DB      1081 TTGTTTGGCACAACAAATCAACGGGACTTTCACAAATGTGTAACAACCTCCGCCATTTG 1140
QY      1141 ACGCAATGGGCGGTAGCGGTGTACGGGTGGAGGTCTATATTAACAGAGCTCGTTTGTG 1200
DB      1141 ACGCAATGGGCGGTAGCGGTGTACGGGTGGAGGTCTATATTAACAGAGCTCGTTTGTG 1200
QY      1201 AACCGTACATGCGCTGAGAGCGCCATTCACGCTGTTTTGACCTCATAGAAAGACACCG 1260
DB      1201 AACCGTACATGCGCTGAGAGCGCCATTCACGCTGTTTTGACCTCATAGAAAGACACCG 1260
QY      1261 GACAAATTCGAGCTGGGTATACCGTGCAGCCACCAAGAGTAAGAAAGATATGAGACT 1320
DB      1261 GACAAATTCGAGCTGGGTATACCGTGCAGCCACCAAGAGTAAGAAAGATATGAGACT 1320
QY      1321 TGTGGAATGGGGGTGGAATGGGGACACATGCTCTTGGGATTTATGATGCTGTAGTG 1380
DB      1321 TGTGGAATGGGGGTGGAATGGGGACACATGCTCTTGGGATTTATGATGCTGTAGTG 1380
QY      1381 CTACAGAAAAATTTGGGTGACCGCTTATTAATGGGGTACCTGTGTGAAGGAACAACA 1440
DB      1381 CTACAGAAAAATTTGGGTGACCGCTTATTAATGGGGTACCTGTGTGAAGGAACAACA 1440
QY      1441 CCACCTATTTTGGCATGCAATGCTAAAGCATATGATACAGAGTACATATGTTTGGG 1500
DB      1441 CCACCTATTTTGGCATGCAATGCTAAAGCATATGATACAGAGTACATATGTTTGGG 1500
QY      1501 CCACCATATGCTGTGTACCCACAGACCCCAACCAAGAAAGTATGTTGTTAAATGTA 1560
DB      1501 CCACCATATGCTGTGTACCCACAGACCCCAACCAAGAAAGTATGTTGTTAAATGTA 1560
QY      1561 CAGAAAAATTTTAAATGATGAAAAATGACATGTGTAAGAACAGATCAAGAGATTAATCA 1620
DB      1561 CAGAAAAATTTTAAATGATGAAAAATGACATGTGTAAGAACAGATCAAGAGATTAATCA 1620
QY      1621 GTTATGGGATCAAAAGCCTAAAGCATGTGTAATTAACCCCACTGTGTTGTTAA 1680
DB      1621 GTTATGGGATCAAAAGCCTAAAGCATGTGTAATTAACCCCACTGTGTTGTTAA 1680
QY      1681 AGTGCATGATTTGAAATGATATCTATATACCAATAGTAGAGGAGATGATTAATGG 1740
DB      1681 AGTGCATGATTTGAAATGATATCTATATACCAATAGTAGAGGAGATGATTAATGG 1740
QY      1741 AGAAGAGAGATTAATAAACTGCTTTCAATATCACACAGCATTAAGAGATTAAGTGC 1800
DB      1741 AGAAGAGAGATTAATAAACTGCTTTCAATATCACACAGCATTAAGAGATTAAGTGC 1800
QY      1801 AGAAGATATGATCTTTTATTAATCTGATATAGTACCAATAGATTAATCAAGCTATA 1860
DB      1801 AGAAGATATGATCTTTTATTAATCTGATATAGTACCAATAGATTAATCAAGCTATA 1860
```

QY	1861	GGTTGATTAAGTGTGTAACACCTCACTCACTATACAGAGGCTGTCCAAAGGTATCCTTTGAGC	1920
Db	1861	GGTTGATTAAGTGTGTAACACCTCACTCACTATACAGAGGCTGTCCAAAGGTATCCTTTGAGC	1920
QY	1921	CAATTCCCATTCATTAATTGTGCCCCGGCTGGTTTGGATTCTTAAATGTAAATATATAGTA	1980
Db	1921	CAATTCCCATTCATTAATTGTGCCCCGGCTGGTTTGGATTCTTAAATGTAAATATATAGTA	1980
QY	1981	CGTTCAATGGAACAGACATGTACAATATGTACAGACAGTACATGTACATGCAATGCA	2040
Db	1981	CGTTCAATGGAACAGACATGTACAATATGTACAGACAGTACATGTACATGCAATGCAATGCA	2040
QY	2041	GGCCAGATATGTACAACTCACTGCTGTTAAATGGCACTGTACAGAAAGAAATGTAGTAA	2100
Db	2041	GGCCAGATATGTACAACTCACTGCTGTTAAATGGCACTGTACAGAAAGAAATGTAGTAA	2100
QY	2101	TTAGATCTGCCAATTTTACAGACAAATGCTTAAACCAATATAGTACAGCTGACACATCTG	2160
Db	2101	TTAGATCTGCCAATTTTACAGACAAATGCTTAAACCAATATAGTACAGCTGACACATCTG	2160
QY	2161	TAGAATTTAATTGTACAGAACCCACAAACAATACAGAAAAAGTATCCGTATCCAGAGG	2220
Db	2161	TAGAATTTAATTGTACAGAACCCACAAACAATACAGAAAAAGTATCCGTATCCAGAGG	2220
QY	2221	GACGAGGAGAGCATTTGTTTACATATAGAAAAAATAGAAATATAGACAGACACATTTGTA	2280
Db	2221	GACGAGGAGAGCATTTGTTTACATATAGAAAAAATAGAAATATAGACAGACACATTTGTA	2280
QY	2281	ACATTAGTAGACAAAAATGGAATGCCACTTTAAACAGATAGCTAGCAATTTAAGAAC	2340
Db	2281	ACATTAGTAGACAAAAATGGAATGCCACTTTAAACAGATAGCTAGCAATTTAAGAAC	2340
QY	2341	AATTTGGAATTAATAAAAACAATATCTTTAAGCATCTCAGAGGGGACCCAGAAATTG	2400
Db	2341	AATTTGGAATTAATAAAAACAATATCTTTAAGCATCTCAGAGGGGACCCAGAAATTG	2400
QY	2401	TAAAGCAGATTTAATTGTGAGAGGGAATTTTCTCTGTAAATTCACACCACTGTGTA	2460
Db	2401	TAAAGCAGATTTAATTGTGAGAGGGAATTTTCTCTGTAAATTCACACCACTGTGTA	2460
QY	2461	ATACTACTTGGTTTAATAGTACTTGTGAGTACTGTAAGGGTCAAAATACACTGAAAGTAGT	2520
Db	2461	ATACTACTTGGTTTAATAGTACTTGTGAGTACTGTAAGGGTCAAAATACACTGAAAGTAGT	2520
QY	2521	ACACAATCACACTCCCTGCGAGATTAATAACAATTTATTAACATGTGCGAGAGGTAGAA	2580
Db	2521	ACACAATCACACTCCCTGCGAGATTAATAACAATTTATTAACATGTGCGAGAGGTAGAA	2580
QY	2581	AAGCAATGTATGCCCTCCCATCGATCGTGACAAATTAATGTTCATCAAAATATTAACGTGGC	2640
Db	2581	AAGCAATGTATGCCCTCCCATCGATCGTGACAAATTAATGTTCATCAAAATATTAACGTGGC	2640
QY	2641	TGCTATTAAACAAGAGATGTGTGTAATTAACAACAATGGGTCGAGATCTTACAGACTGGAG	2700
Db	2641	TGCTATTAAACAAGAGATGTGTGTAATTAACAACAATGGGTCGAGATCTTACAGACTGGAG	2700
QY	2701	GAGCGATATGAGGAGCAATTTGGAGAAAGTATATTAATAATTAAGGTGTAAAAATTG	2760
Db	2701	GAGCGATATGAGGAGCAATTTGGAGAAAGTATATTAATAATTAAGGTGTAAAAATTG	2760
QY	2761	AACCAATTAGAGTGTGACCCACCAAGCAAGCAAGAAAGAAAGAGAGTGTGCGAGAGAAAAAGAG	2820
Db	2761	AACCAATTAGAGTGTGACCCACCAAGCAAGCAAGAAAGAAAGAGAGTGTGCGAGAGAAAAAGAG	2820
QY	2821	CAGTGGGAATTAAGAGCTTTGTTCTCCTTGCTTGGAGCAGCAGAGAGGACACTATGGCT	2880
Db	2821	CAGTGGGAATTAAGAGCTTTGTTCTCCTTGCTTGGAGCAGCAGAGAGGACACTATGGCT	2880
QY	2881	GCAGGTCAATGACCTGACGCTTACAGGCCACAGACATTAATGTCTGATATATGTGACGAGC	2940
Db	2881	GCAGGTCAATGACCTGACGCTTACAGGCCACAGACATTAATGTCTGATATATGTGACGAGC	2940
QY	4021	ACAAACCAACACAGAAATGACAGTGAAGAAAAATGCTTATTTGTAATTTGTAATTTGTAAT	4080
Db	4021	ACAAACCAACACAGAAATGACAGTGAAGAAAAATGCTTATTTGTAATTTGTAATTTGTAAT	4080

```
|||||
Db 4021 ACAACCCACATGAAATGACAGTGAATAAATGCTTATTTGTGAATTTGTGATGCTAT 4080
Qy 4081 TGCCTTATTTGTATACATATTAAGCTGCAATTAACAGATTAACAACAACATTTGCTATCA 4140
Db 4081 TGCCTTATTTGTATTAACATTAATAGCTGCAATTAACAGATTAACAACAACATTTGCTATCA 4140
Qy 4141 TTTTATGTTTCAGGTTTCAGGGGAGGTGTGGAGGTTTAAAGCAAGTAAACCTCTA 4200
Db 4141 TTTTATGTTTCAGGTTTCAGGGGAGGTGTGGAGGTTTAAAGCAAGTAAACCTCTA 4200
Qy 4201 CAAATGTGTATGCTGATTAATGATTCCTGCTCGCGCTTTCGGTGATGACGTGAAGC 4260
Db 4201 CAAATGTGTATGCTGATTAATGATTCCTGCTCGCGCTTTCGGTGATGACGTGAAGC 4260
Qy 4261 CTCGACACATGACAGCTCCCGGAGAGGTCACAGCTTGTCTGTAAAGCGGATGCCGGAGC 4320
Db 4261 CTCGACACATGACAGCTCCCGGAGAGGTCACAGCTTGTCTGTAAAGCGGATGCCGGAGC 4320
Qy 4321 AGACAAGCCCTCAGGGGCGCTCAGCGGGTGTGGCGGGTGTGGGGCGCAGCCATGACC 4380
Db 4321 AGACAAGCCCTCAGGGGCGCTCAGCGGGTGTGGCGGGTGTGGGGCGCAGCCATGACC 4380
Qy 4381 CAGTCAGCTAGCGATAGCGGAGTGTATCTGCTTAACTATGCGGCATCAGAGCAGATG 4440
Db 4381 CAGTCAGCTAGCGATAGCGGAGTGTATCTGCTTAACTATGCGGCATCAGAGCAGATG 4440
Qy 4441 TACTGAGAGTGCACCATATGTCGGGCGCGCTTGTGGGGTTCATAGGCTCCGGCCC 4500
Db 4441 TACTGAGAGTGCACCATATGTCGGGCGCGCTTGTGGGGTTCATAGGCTCCGGCCC 4500
Qy 4501 CCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAACCAGAGACTA 4560
Db 4501 CCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAACCAGAGACTA 4560
Qy 4561 TAAAGATACAGAGGTTTCCCGCTGGAAGTCCCTCGGCGCTCTCCTGTTCCGACCTG 4620
Db 4561 TAAAGATACAGAGGTTTCCCGCTGGAAGTCCCTCGGCGCTCTCCTGTTCCGACCTG 4620
Qy 4621 CCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAAGCTGGCGCTTTCATATG 4680
Db 4621 CCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAAGCTGGCGCTTTCATATG 4680
Qy 4681 TCACGCTGTAGATCTCAGTTGCGTGTAGTGTGCTGCTCAAGCTGGGCTGTGTGAC 4740
Db 4681 TCACGCTGTAGATCTCAGTTGCGTGTAGTGTGCTGCTCAAGCTGGGCTGTGTGAC 4740
Qy 4741 GAACCCCCCTTACGCCCCGAGCCGCTTATCCGCTAATCTGCTTGAATCCAAC 4800
Db 4741 GAACCCCCCTTACGCCCCGAGCCGCTTATCCGCTAATCTGCTTGAATCCAAC 4800
Qy 4801 CCGGTAAGACACGACTTATCCGCTGAGCAGCAGCAGTGTAAACAGGATTAAGCAGCG 4860
Db 4801 CCGGTAAGACACGACTTATCCGCTGAGCAGCAGCAGTGTAAACAGGATTAAGCAGCG 4860
Qy 4861 AGGATGTAGGCGGTGTACAGAGTCTTGAAGTGTGGCCTAATCTACGCGCTACACTAGA 4920
Db 4861 AGGATGTAGGCGGTGTACAGAGTCTTGAAGTGTGGCCTAATCTACGCGCTACACTAGA 4920
Qy 4921 AGGACAGTATTTGCTATCTGCGCTCTCTGAAAGCAGTTACTCTCGGAAAAAGAGTGGT 4980
Db 4921 AGGACAGTATTTGCTATCTGCGCTCTCTGAAAGCAGTTACTCTCGGAAAAAGAGTGGT 4980
Qy 4981 AGCTTGTATCCGGCAAAACAAACGACGCGGTGTAGCGGTGTTTTTTTGTGAAGCAG 5040
Db 4981 AGCTTGTATCCGGCAAAACAAACGACGCGGTGTAGCGGTGTTTTTTTGTGAAGCAG 5040
Qy 5041 CAGATTACGCGCAGAAAAAGATCTCAGAAAGATCTTGTGATCTTTCTACGGGGCT 5100
Db 5041 CAGATTACGCGCAGAAAAAGATCTCAGAAAGATCTTGTGATCTTTCTACGGGGCT 5100
Qy 5101 GACGCTAGTGAAGCAAACTACGTTAAGGATTTTGTGATGAGATTATCAAAAAG 5160
Db 5101 GACGCTAGTGAAGCAAACTACGTTAAGGATTTTGTGATGAGATTATCAAAAAG 5160

|||||
Db 5101 GACGCTAGTGAAGCAAACTACGTTAAGGATTTTGTGATGAGATTATCAAAAAG 5160
Qy 5161 ATCTCACCTGATACCTTTTAAATTAAGTTTAAATCAATCAATCAATATAT 5220
Db 5161 ATCTCACCTGATACCTTTTAAATTAAGTTTAAATCAATCAATCAATATAT 5220
Qy 5221 GAGTAAACTGTGTGACAGTTTACCAATGCTTAATCAAGTACAGCCTATCTCAGCATC 5280
Db 5221 GAGTAAACTGTGTGACAGTTTACCAATGCTTAATCAAGTACAGCCTATCTCAGCATC 5280
Qy 5281 TGTCTATTTGCTTATCATATGTTGCTGACTCCCGCTCGTGTAGATACATACATACG 5340
Db 5281 TGTCTATTTGCTTATCATATGTTGCTGACTCCCGCTCGTGTAGATACATACATACG 5340
Qy 5341 GAGGCGTTACCATGTGGCCCGACAGTGTGCATGATGTACCGGAGACCCAGCTCACGGCT 5400
Db 5341 GAGGCGTTACCATGTGGCCCGACAGTGTGCATGATGTACCGGAGACCCAGCTCACGGCT 5400
Qy 5401 CCAGATTTATCAGCAATTAACACAGCCGGAAGGGCCGACAGCCAGAAAGTGTCTGCA 5460
Db 5401 CCAGATTTATCAGCAATTAACACAGCCGGAAGGGCCGACAGCCAGAAAGTGTCTGCA 5460
Qy 5461 ACTTTATCCGCTTCATATCAGTCTATTAATTTGTCGGGGAAGCTAAGTAAGTATCG 5520
Db 5461 ACTTTATCCGCTTCATATCAGTCTATTAATTTGTCGGGGAAGCTAAGTAAGTATCG 5520
Qy 5521 CCAGTTAATGATTTGGCCCAAGGTTGTCGCTCTCTACAGCATCTGTGTGTCAGCTCG 5580
Db 5521 CCAGTTAATGATTTGGCCCAAGGTTGTCGCTCTCTACAGCATCTGTGTGTCAGCTCG 5580
Qy 5581 TCGTTTGTATGCTTCATTCAGCTCCGGTCCCAAGCATCAAGGCGATTAATGATTC 5640
Db 5581 TCGTTTGTATGCTTCATTCAGCTCCGGTCCCAAGCATCAAGGCGATTAATGATTC 5640
Qy 5641 CCCATGTTGTGCAAAAAAGGGGTAGCTCCTCGGCTCCGATCGTTGTGAGAAATAG 5700
Db 5641 CCCATGTTGTGCAAAAAAGGGGTAGCTCCTCGGCTCCGATCGTTGTGAGAAATAG 5700
Qy 5701 TTGGCCCGAGTGTATCATCATGTTATGTGACACACATGCAATATCTCTTACGTATG 5760
Db 5701 TTGGCCCGAGTGTATCATCATGTTATGTGACACACATGCAATATCTCTTACGTATG 5760
Qy 5761 CCATCCGTAAGATGCTTTTCTGTGACGTGTGATCTCAACCAAGTCAATCTGGAATAG 5820
Db 5761 CCATCCGTAAGATGCTTTTCTGTGACGTGTGATCTCAACCAAGTCAATCTGGAATAG 5820
Qy 5821 TGTATGGGGGACCGAGTTGCTGTCGGCCGCGCATTAAGGGATTAATACCGGCGCAT 5880
Db 5821 TGTATGGGGGACCGAGTTGCTGTCGGCCGCGCATTAAGGGATTAATACCGGCGCAT 5880
Qy 5881 AGCAGAACTTTAAAGTGTATCATGTTGAAGAGTTCTTCGGGGCAAAACTCTCAAG 5940
Db 5881 AGCAGAACTTTAAAGTGTATCATGTTGAAGAGTTCTTCGGGGCAAAACTCTCAAG 5940
Qy 5941 ATCTTACCGCTGTTGAGATCCAGTTGATGTAAACCACTGTCGACCCAACTGATCTTCA 6000
Db 5941 ATCTTACCGCTGTTGAGATCCAGTTGATGTAAACCACTGTCGACCCAACTGATCTTCA 6000
Qy 6001 GCATCTTTTACTTTCACACAGCTTTCGGGTGAGCAAAACAGAAAGCAAAATGCCGA 6060
Db 6001 GCATCTTTTACTTTCACACAGCTTTCGGGTGAGCAAAACAGAAAGCAAAATGCCGA 6060
Qy 6061 AAAAAGGAATTAAGGGGACACAGGAATGTTGAATTAATCTATCTCTTTTCAATAT 6120
Db 6061 AAAAAGGAATTAAGGGGACACAGGAATGTTGAATTAATCTATCTCTTTTCAATAT 6120
Qy 6121 TATTAAGCATTTATCAAGGTTATTTGTCATGAGCGGATACATATTTGAATGATTTAG 6180
Db 6121 TATTAAGCATTTATCAAGGTTATTTGTCATGAGCGGATACATATTTGAATGATTTAG 6180
Qy 6181 AAAAATTAACAATAGGGGTTCCGCGCATATTTCCCGAAAAAGTGCAC 6229
Db 6181 AAAAATTAACAATAGGGGTTCCGCGCATATTTCCCGAAAAAGTGCAC 6229
```


RESULT 2
AR067813 LOCUS AR067813 9709 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 5 from patent US 5851813.
ACCESSION AR067813
VERSION AR067813.1 GI:5999035
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 9709)
AUTHORS Desrosiers,R.C.
TITLE Primate lentivirus antigenic compositions
JOURNAL Patent: US 5851813-A 5 22-DEC-1998;
FEATURES
Source location/Qualifiers
1..9709
BASE COUNT 3421 a 1756 c 2366 g 2166 t
ORIGIN
Query Match 43.0%; Score 2676.2; DB 6; Length 9709;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1291 CCATGAGTGAAGAGAGATATCAGCACTGTGGAGATGGGGTGAATGGGGACCA 1350
DB 6219 CATGAGAGTGAAGAGAGATATCAGCACTGTGGAGATGGGGTGAATGGGGACCA 6278
QY 1351 TGCCTCTGGGATATGATGATCTGTAGTCTACAGAAAAATTTGGGTCACCGTCTATT 1410
DB 6279 TGCCTCTGGGATATGATGATCTGTAGTCTACAGAAAAATTTGGGTCACCGTCTATT 6338
QY 1411 ATGGGATACCTGTGTGGAGAGACACACCACCTCTATTGTCATCAGATGCTAAAG 1470
DB 6339 ATGGGATACCTGTGTGGAGAGACACACCACCTCTATTGTCATCAGATGCTAAAG 6398
QY 1471 CATATGATACAGAGGTATCATATGTTGGGCCACACATGCGTGTACCCACAGCCCA 1530
DB 6399 CATATGATACAGAGGTATCATATGTTGGGCCACACATGCGTGTACCCACAGCCCA 6458
QY 1531 ACCCACAAGAGTATGTTGTTAAATGTGACAGAAAAATTTTAACTGTGAAAAATGACA 1590
DB 6459 ACCCACAAGAGTATGTTGTTAAATGTGACAGAAAAATTTTAACTGTGAAAAATGACA 6518
QY 1591 TGGTAGAACAAGATGATGATATATCAGTTATGGATCAAAAAGCCTAAAGCCATGTG 1650
DB 6519 TGGTAGAACAAGATGATGATATATCAGTTATGGATCAAAAAGCCTAAAGCCATGTG 6578
QY 1651 TAAATTAACCCACTGCTGTAGTTAAAGTGCACATGTTGAAGAAATGATACTATA 1710
DB 6579 TAAATTAACCCACTGCTGTAGTTAAAGTGCACATGTTGAAGAAATGATACTATA 6638
QY 1711 CCAATAGTAGAGGGGAGATGATATGAGAAAAAGAGATAAAAAAGTCTCTTCA 1770
DB 6639 CCAATAGTAGAGGGGAGATGATATGAGAAAAAGAGATAAAAAAGTCTCTTCA 6698
QY 1771 ATATCAGACAAGATGATGATAGAGATAGGTCAGAAAAAGATATGATCTTTTAAACTTG 1830
DB 6699 ATATCAGACAAGATGATGATAGAGATAGGTCAGAAAAAGATATGATCTTTTAAACTTG 6758
QY 1831 ATATAGTACCAATGATATATACAGCTATAGTTGATAGTTGTAACCTCAGTCATTA 1890
DB 6759 ATATAGTACCAATGATATATACAGCTATAGTTGATAGTTGTAACCTCAGTCATTA 6818
QY 1891 CACAGAGCTGTGCAAAAGTATCTTTGAGCAATCCATATCATTTATTTGGCCCGGCTG 1950
DB 6819 CACAGAGCTGTGCAAAAGTATCTTTGAGCAATCCATATCATTTATTTGGCCCGGCTG 6878
QY 1951 GTTTTGCATTTCTAAATATGATATAGAGTTCAATGAGAACAGACCATGTACAAATG 2010
DB 6879 GTTTTGCATTTCTAAATATGATATAGAGTTCAATGAGAACAGACCATGTACAAATG 6938

QY 2011 TCAGCAGTACCAATGTACACATGGAATCAGGCCAGTAGATCAACTCAACTGCTTTAA 2070
DB 6939 TCAGCAGTACCAATGTACACATGGAATCAGGCCAGTAGATCAACTCAACTGCTTTAA 6998
QY 2071 ATGGCAGTCTAGCAGAAAGATGATGATTAATGATCTGCCAATTTTCACAGACATGCTA 2130
DB 6999 ATGGCAGTCTAGCAGAAAGATGATGATTAATGATCTGCCAATTTTCACAGACATGCTA 7058
QY 2131 AAACCATATATGATCAGTGAACACATCTGTGAATTAATTTGTAAGACCCACAACA 2190
DB 7059 AAACCATATATGATCAGTGAACACATCTGTGAATTAATTTGTAAGACCCACAACA 7118
QY 2191 ATACAGAAAAAGATATCCGTATCCAGAGGGGACCGAGAGCATTTGTTCAATAGGA 2250
DB 7119 ATACAGAAAAAGATATCCGTATCCAGAGGGGACCGAGAGCATTTTTCATAGGA 7178
QY 2251 AAATAGAAATATGAGACACACATTTGTAACTATAGAGCAAAATGGAATGCCACTT 2310
DB 7179 AAATAGAAATATGAGACACACATTTGTAACTATAGAGCAAAATGGAATGCCACTT 7238
QY 2311 TAAACAGATATGCTAGCAATTTAGAGAACATTTTGGAAATATATTAACATATCTTTA 2370
DB 7239 TAAACAGATATGCTAGCAATTTAGAGAACATTTTGGAAATATATTAACATATCTTTA 7298
QY 2371 AGCAATCTCAGAGAGGGAGCCAGAAATTTGTAAACAGATTTTAAATGTGAGGGGAAT 2430
DB 7299 AGCAATCTCAGAGAGGGAGCCAGAAATTTGTAAACAGATTTTAAATGTGAGGGGAAT 7358
QY 2431 TTTTCTACTGTAATTTCAACACACATGTTTAATAGTACTGTTTAATAGTACTTGGAGTA 2490
DB 7359 TTTTCTACTGTAATTTCAACACACATGTTTAATAGTACTGTTTAATAGTACTTGGAGTA 7418
QY 2491 CTGAAGGTCATATTAACCTGAGAGAAATGACACATACACTCCCATGCAAGATTAAC 2550
DB 7419 CTGAAGGTCATATTAACCTGAGAGAAATGACACATACACTCCCATGCAAGATTAAC 7478
QY 2551 AATTATTAACATGAGGAGAGAGAGAAAGCAATGATGCCCCCTCCCATCACTGGAG 2610
DB 7479 AATTATTAACATGAGGAGAGAGAGAAAGCAATGATGCCCCCTCCCATCACTGGAG 7538
QY 2611 AATTATTAACATGAGGAGAGAGAGAAAGCAATGATGCCCCCTCCCATCACTGGAG 2670
DB 7539 AATTATTAACATGAGGAGAGAGAGAAAGCAATGATGCCCCCTCCCATCACTGGAG 7598
QY 2671 ACAATGGTCCGAGATCTTCAGACCTGAGAGAGCGATATGAGGACATTTGGAGATG 2730
DB 7599 ACAATGGTCCGAGATCTTCAGACCTGAGAGAGCGATATGAGGACATTTGGAGATG 7658
QY 2731 AATTATTAACATGAGGAGAGAGAGAAAGCAATGATGCCCCCTCCCATCACTGGAG 2790
DB 7659 AATTATTAACATGAGGAGAGAGAGAAAGCAATGATGCCCCCTCCCATCACTGGAG 7718
QY 2791 AGAGAAGAGTGTGAG 2850
DB 7719 AGAGAAGAGTGTGAG 7778
QY 2851 TCTTGGAG 2910
DB 7779 TCTTGGAG 7838
QY 2911 GACAAATTTGCTGATATAGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2970
DB 7839 GACAAATTTGCTGATATAGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7898
QY 2971 AACAGCATCTGTGCAATCTACAGAGTGGGCAATCAACAGCTCCAGCAGAAATCTGG 3030
DB 7899 AACAGCATCTGTGCAATCTACAGAGTGGGCAATCAACAGCTCCAGCAGAAATCTGG 7958
QY 3031 CTGTGAGAAAGATACCTAAAGAGATCAACAGCTCTGGGAGATTTGGGTTGCTGTGAAAC 3090
DB 7959 CTGTGAGAAAGATACCTAAAGAGATCAACAGCTCTGGGAGATTTGGGTTGCTGTGAAAC 8018

3091 TCATTGGACCACTGCTGTGCTTGGATGCTAGTTGGAGTAATTAATCTCTGACACAGA 3150
|||||
8019 TCATTGGACCACTGCTGTGCTTGGATGCTAGTTGGAGTAATTAATCTCTGACACAGA 8078
|||||
3151 TTTGGATTAACATGACCTGGATGGAGTGGACAGAGAAATTAACAAATTAACAAAGCTTAA 3210
|||||
8079 TTTGGAAATTAACATGACCTGGATGGAGTGGACAGAGAAATTAACAAATTAACAAAGCTTAA 8138
|||||
3211 TACACTCCTTAATTAAGAAATCGCAAAACAGCAAGAAAGATGAACAAGAAATTAATGG 3270
|||||
8139 TACACTCCTTAATTAAGAAATCGCAAAACAGCAAGAAAGATGAACAAGAAATTAATGG 8198
|||||
3271 AATTAGATTAATGGGCAAGTTTGGAAATGTTTAACATAACAAATTGCTGTGTATA 3330
|||||
8199 AATTAGATTAATGGGCAAGTTTGGAAATGTTTAACATAACAAATTGCTGTGTATA 8258
|||||
3331 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3390
|||||
8259 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8318
|||||
3391 TTTCTAATAGTAATAGATTAAGGAGGATATTCACCAATTAATTCGTTTCAGACCCACCTCC 3450
|||||
8319 TTTCTAATAGTAATAGATTAAGGAGGATATTCACCAATTAATTCGTTTCAGACCCACCTCC 8378
|||||
3451 CAATCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAAGTGGAGAGAGACA 3510
|||||
8379 CAATCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAAGTGGAGAGAGACA 8438
|||||
3511 GAGACAGATTCATTCGATTAAGTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGAT 3570
|||||
8439 GAGACAGATTCATTCGATTAAGTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGAT 8498
|||||
3571 GCTGTGCTCTTCAGCTACACCGCTTGAAGAGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGAT 3630
|||||
8499 GCTGTGCTCTTCAGCTACACCGCTTGAAGAGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGAT 8558
|||||
3631 TGGAACTTCTGGAGACGAGGGGGTGGGAAGCCCTCAATTAATTTGGTGAATCTCCTACAGT 3690
|||||
8559 TGGAACTTCTGGAGACGAGGGGGTGGGAAGCCCTCAATTAATTTGGTGAATCTCCTACAGT 8618
|||||
3691 ATTGAGTACAGAACTAAGAAATAGTGTCTTAATCTGCTCAATGCTCCACAGCCATTAAGCAG 3750
|||||
8619 ATTGAGTACAGAACTAAGAAATAGTGTCTTAATCTGCTCAATGCTCCACAGCCATTAAGCAG 8678
|||||
3751 TAGCTGAGGGGACAGATAGGGTTATAGAAGTATTACAAGCAGCTTAATAGACCTAATTCGCC 3810
|||||
8679 TAGCTGAGGGGACAGATAGGGTTATAGAAGTATTACAAGCAGCTTAATAGACCTAATTCGCC 8738
|||||
3811 ACATACCTAGAAAGAAATAAGACAGAGGCTTGGAAAGATTTGCTTAAGATGGGTGGCAAG 3870
|||||
8739 ACATACCTAGAAAGAAATAAGACAGAGGCTTGGAAAGATTTGCTTAAGATGGGTGGCAAG 8798
|||||
3871 TGTGTCAAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3930
|||||
8799 TGTGTCAAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8858
|||||
3931 CCAGCAGCAGATGGGGTGGAGAGCAGTATCTCGAGATCTAGA 3971
|||||
8859 CCAGCAGCAGATGGGGTGGAGAGCAGTATCTCGAGATCTAGA 8899
|||||

RESULT 3
AR224437 9709 bp DNA linear PAT 26-SEP-2002
LOCUS AR224437
DEFINITION Sequence 15 from patent US 6440730.
ACCESSION AR224437
VERSION AR224437.1 GI:23333237
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9709)
AUTHORS Von Laer, M.-D. and Beyer, W.

TITLE Retroviral hybrid vectors pseudotyped with LCMV
JOURNAL Patent: US 6440730-A 15 27-AUG-2002;
FEATURES location/qualifiers
SOURCE 1. 9709 /organism="unknown"
BASE COUNT 3421 a 1756 c 2366 g 2166 t
ORIGIN
Query Match 43.0%; Score 2676.2; DB 6; Length 9709;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
1291 CCATGAGAGTGAAGAGAGATATCAGCAGCTTGTGAGATGGGGTGGAAATGGGGACCA 1350
6219 CAATGAGAGTGAAGAGAGATATCAGCAGCTTGTGAGATGGGGTGGAAATGGGGACCA 6278
1351 TGTCTCTTGGGATTTGTATGATCTGTAGTGTCTACAGAAATTTGGGTCACTCTAAT 1410
6279 TGTCTCTTGGGATTTGTATGATCTGTAGTGTCTACAGAAATTTGGGTCACTCTAAT 6338
1411 ATGGGGTACCTGTGTGGAGAGCAACACACCTCTATTTTGTGATCAGATGCTAAG 1470
6339 ATGGGGTACCTGTGTGGAGAGCAACACACCTCTATTTTGTGATCAGATGCTAAG 6398
1471 CATATGATACAGAGTACATATGTTTGGGACACATGCTGTGTACCCACAGACCCA 1530
6399 CATATGATACAGAGTACATATGTTTGGGACACATGCTGTGTACCCACAGACCCA 6458
1531 ACCCACAAGAGTATGTTTGGTAAATGTGACAGAAATTTTACATGTGGAAATATACA 1590
6459 ACCCACAAGAGTATGTTTGGTAAATGTGACAGAAATTTTACATGTGGAAATATACA 6518
1591 TGTGAGAAACAGATACAGAGATATATCAAGTTTATGGATCAAAAGCCTAAAGCATGTG 1650
6519 TGTGAGAAACAGATACAGAGATATATCAAGTTTATGGATCAAAAGCCTAAAGCATGTG 6578
1651 TAAATTAACCCACTCTGTGTATGTTTAAAGTCACTGATTTGAGACATGATATATA 1710
6579 TAAATTAACCCACTCTGTGTATGTTTAAAGTCACTGATTTGAGACATGATATATA 6638
1711 CCATATGATAGAGGGGAGATGATTAATGGAAGAAAGAGATTAATAAATGCTCTTCA 1770
6639 CCATATGATAGAGGGGAGATGATTAATGGAAGAAAGAGATTAATAAATGCTCTTCA 6698
1771 ATATCAGACAAAGCATTAAGATTAAGTGAAGAAAGATTAATGATCTTTTAAACTTG 1830
6699 ATATCAGACAAAGCATTAAGATTAAGTGAAGAAAGATTAATGATCTTTTAAACTTG 6758
1831 ATATGATACCAATTAAGATTAACAGCTATAGTGTATGATTAAGTGTATACACCTCACTTA 1890
6759 ATATGATACCAATTAAGATTAACAGCTATAGTGTATGATTAAGTGTATACACCTCACTTA 6818
1891 CACAGGCTGTCCCAAGGATATCTTTGAGCAATTTCCCATATCAATTAATTTGGCCCGCTG 1950
6819 CACAGGCTGTCCCAAGGATATCTTTGAGCAATTTCCCATATCAATTAATTTGGCCCGCTG 6878
1951 GTTTTGGGATTTTAAATTAATTAATTAAGAGTTCAATGAGACAGACATGATACAATG 2010
6879 GTTTTGGGATTTTAAATTAATTAATTAAGAGTTCAATGAGACAGACATGATACAATG 6938
2011 TCAGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2070
6939 TCAGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6998
2071 ATGGCAGCTAGCAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2130
6999 ATGGCAGCTAGCAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7058
2131 AAACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2190
7059 AAACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7118
2191 ATACAGAAAAAGTATCCGTATCCAGAGGGGACAGGAGACATTTTGTATCAATTAAGAA 2250

Db	7119	ATACAGAAAAAGTATCCGTATCCAGAGGGGACCAAGGAGCAATTGTTCAAATAGGAA	7178
Qy	2251	AAATAGGAAATATAGACAAAGCAATTTTAACATATAGAGACAAATAGGATGCCATT	2310
Db	7179	AAATAGGAAATATAGACAAAGCAATTTTAACATATAGAGACAAATAGGATGCCATT	7238
Qy	2311	TAAACAGATATGCTAGCAAAATTTAAGACAACATTTGGAAATATTAACAAATTAATCTTTA	2370
Db	7239	TAAACAGATATGCTAGCAAAATTTAAGACAACATTTGGAAATATTAACAAATTAATCTTTA	7298
Qy	2371	AGCAATCCCTCGAGGGGGACCCACAATAATTTGACGCAAGTTTAATTTGTGGAGGGAGT	2438
Db	7299	AGCAATCCCTCGAGGGGGACCCACAATAATTTGACGCAAGTTTAATTTGTGGAGGGAGT	7358
Qy	2431	TTTTCTACTGTAATTCACACAACTGTTTAATAGTACTTGGTTTAATATGATCTTGAGATA	2490
Db	7359	TTTTCTACTGTAATTCACACAACTGTTTAATAGTACTTGGTTTAATATGATCTTGAGATA	7418
Qy	2491	CTGAAGGTCCTCAATTAACCTGTAAGAGAGTACACAACTCCCATGCAAGATTAAC	2550
Db	7419	CTGAAGGTCCTCAATTAACCTGTAAGAGAGTACACAACTCCCATGCAAGATTAAC	7478
Qy	2551	AATTTATTAACATGTGGCAGGAATAGGAAAAGCAATGTATGCCCTCCATCACTGAC	2610
Db	7479	AATTTATTAACATGTGGCAGGAATAGGAAAAGCAATGTATGCCCTCCATCACTGAC	7538
Qy	2611	AAATTAGATGTCATCAAAATTTACTGGGCTGCTATTACAAAGAGATGGTGGTATACA	2670
Db	7539	AAATTAGATGTCATCAAAATTTACTGGGCTGCTATTACAAAGAGATGGTGGTATACA	7598
Qy	2671	ACAATGGGTCGAGATCTTCAGACCTCGAGAGGCGATATAGAGGACAAATTTGGAGAA	2730
Db	7599	ACAATGGGTCGAGATCTTCAGACCTCGAGAGGCGATATAGAGGACAAATTTGGAGAA	7658
Qy	2731	AATTTATTAATATTAAGTACTAAAAATTTGAACATTAGAGTATAGCACCCACCAAGCAA	2790
Db	7659	AATTTATTAATATTAAGTACTAAAAATTTGAACATTAGAGTATAGCACCCACCAAGCAA	7718
Qy	2791	AGAGACAGTGTCTCAGAGAGAAAAAGACAGTGGGAAATAGAGACTTTGTCTCTGGCT	2850
Db	7719	AGAGACAGTGTCTCAGAGAGAAAAAGACAGTGGGAAATAGAGACTTTGTCTCTGGCT	7778
Qy	2851	TCTTGGGAGCAGCAGGAAGCACTATGGGCTTCACAGTCAATACCTGACGGTACAGGCCA	2910
Db	7779	TCTTGGGAGCAGCAGGAAGCACTATGGGCTTCACAGTCAATACCTGACGGTACAGGCCA	7838
Qy	2911	GACATTTATTTGTCTGATATATAGTGCAGCAGCAAGAAATTTTCTGAGGGCTATGAGCGC	2970
Db	7839	GACATTTATTTGTCTGATATATAGTGCAGCAGCAAGAAATTTTCTGAGGGCTATGAGCGC	7898
Qy	2971	AACGACATCTGTTCACACTCACTCTGAGGGGATCCAAACAGCTCCAGGCAAGATCTCGG	3030
Db	7899	AACGACATCTGTTCACACTCACTCTGAGGGGATCCAAACAGCTCCAGGCAAGATCTCGG	7958
Qy	3031	CTGTGGAAGAATTAACCTTAAGAGTCAACAGCTCTGGGGATTTGGGTTGCTCTGGAAC	3090
Db	7959	CTGTGGAAGAATTAACCTTAAGAGTCAACAGCTCTGGGGATTTGGGTTGCTCTGGAAC	8018
Qy	3091	TCATTTGCACCACTGCTGTGCCCTTGGAAATGCTAGTTGGAGTAATTAATCTCTGAAACA	3150
Db	8019	TCATTTGCACCACTGCTGTGCCCTTGGAAATGCTAGTTGGAGTAATTAATCTCTGAAACA	8078
Qy	3151	TTTGGAAATTAACATCACTCGATATGAGATGAGACAGAGAAATTAACATTTACCAAGCTTAA	3210
Db	8079	TTTGGAAATTAACATCACTCGATATGAGATGAGACAGAGAAATTAACATTTACCAAGCTTAA	8138
Qy	3211	TACACTCTCTTAATTTGAAGAATTCGAAAAACCGACAAGAAAAAGATGAACAATTAATTTGG	3270
Db	8139	TACACTCTCTTAATTTGAAGAATTCGAAAAACCGACAAGAAAAAGATGAACAATTAATTTGG	8198
Qy	3271	AATTTAGATTAATGGGCAAGTTTGTGGAATGCTTTAAACATTAACAAATTTGCTGTGGTATA	3330

Db	8199	AATTAGATPAAATGGCCAACTTTGTGCAATGGTGTAAACATAACAAATTTGGCTGTGATATA	8258		
QY	3331	TAAATTTATTCATATATGATAGTAGAGGCTTGTAAGTTTAAATAGTTTTCCTGTAC	3390		
Db	8259	TAAATTTATTCATATATGATAGTAGAGGCTTGTAAGTTTAAATAGTTTTCCTGTAC	8318		
QY	3391	TTTCTATATAGTATAGAGTTAGCGAGGATATTCACATTAATTCGTTTCAGACCACTTC	3450		
Db	8319	TTTCTATATAGTATAGAGTTAGCGAGGATATTCACATTAATTCGTTTCAGACCACTTC	8378		
QY	3451	CAATCCGAGGGAGCCCGACAGGCCCGCAAGAAATATGAAGAAGAGTGTGAGAGAGACA	3510		
Db	8379	CAATCCGAGGGAGCCCGACAGGCCCGCAAGAAATATGAAGAAGAGTGTGAGAGAGACA	8438		
QY	3511	GAGACAGATTCATTCGATTAGTATGAAGCGATCCTTAGCCTTAATTCGTGGAGCATCTGCGGA	3570		
Db	8439	GAGACAGATTCATTCGATTAGTATGAAGCGATCCTTAGCCTTAATTCGTGGAGCATCTGCGGA	8498		
QY	3571	GCCTGTGCTCTTCAGCTTACCAACCGCTTGAGAGACTTACTCTGATTTGAACAAGAGATTG	3630		
Db	8499	GCCTGTGCTCTTCAGCTTACCAACCGCTTGAGAGACTTACTCTGATTTGAACAAGAGATTG	8558		
QY	3631	TGGAACCTTCTGGGAGCGAGGGGGTGTGGAAGCCCTCAATATTTGGTGGAAATCTCTACAGT	3690		
Db	8559	TGGAACCTTCTGGGAGCGAGGGGGTGTGGAAGCCCTCAATATTTGGTGGAAATCTCTACAGT	8618		
QY	3691	ATTGAGATCAGGAACCTTAAGAATAGTGTCTTAACCTTCCTCAATGCCACAGCATATGACG	3750		
Db	8619	ATTGAGATCAGGAACCTTAAGAATAGTGTCTTAACCTTCCTCAATGCCACAGCATATGACG	8678		
QY	3751	TAGCTGAGGGGACAGATATGAGGTTATAGAAGATTATACAAGACGCTTATAGACTTATTCGCC	3810		
Db	8679	TAGCTGAGGGGACAGATATGAGGTTATAGAAGATTATACAAGACGCTTATAGACTTATTCGCC	8738		
QY	3811	ACATACCTTGAAGAATTAAGACAGGGCTTGGAAGAGATTTTCTTATAGATGGGTGGCAAG	3870		
Db	8739	ACATACCTTGAAGAATTAAGACAGGGCTTGGAAGAGATTTTCTTATAGATGGGTGGCAAG	8798		
QY	3871	TGGTCAAAAAGTAGTGATTTGGATGGCTGCTCTGTAGAGGGAAGAATGAGACGAGCTGAG	3930		
Db	8799	TGGTCAAAAAGTAGTGATTTGGATGGCTGCTCTGTAGAGGGAAGAATGAGACGAGCTGAG	8858		
QY	3931	CCAGCAGCAGATGGGGTGGAGACAGTATCTCGAGATCTAGA 3971			
Db	8859	CCAGCAGCAGATGGGGTGGAGACAGTATCTCGAGACCTAGA 8899			
RESULT 4	AX032749	9709 bp	DNA	Linear	PAT 21-SEP-2000
LOCUS	AX032749	Sequence 1 from Patent WO0047223.			
DEFINITION	AX032749	AX032749.1 GI:10279727			
ACCESSION	AX032749				
VERSION	AX032749.1				
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
BASE COUNT					
ORIGIN					
Query Match	43.0%	Score 2676.2	DB 6	Length 9709	
Best Local Similarity	99.9%	Pred. No. 0			

Matches 2678: Conservative 0; Mismatches 3; Indels 0; Gaps 0;				
Qy	1291	CCATGAGAGTGAAGAGAGATATCAGCACTGTGTGAGATGGGGGTGGAATGGGGCA	1350	
Db	6219	CAATGAGAGTGAAGAGAGATATCAGCACTGTGTGAGATGGGGGTGGAATGGGGCA	6278	
Qy	1351	TGCTCCTTGGGATATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTAT	1410	
Db	6279	TGCTCCTTGGGATATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTAT	6338	
Qy	1411	ATGGGGTACCTGTGTGAGAGAACCAACCACTATTTTGTGCAATTCAGATCTTAAG	1470	
Db	6339	ATGGGGTACCTGTGTGAGAGAACCAACCACTATTTTGTGCAATTCAGATCTTAAG	6398	
Qy	1471	CATATGATACGAGAGTACATATGTATGTATGTATGTATGTATGTATGTATGTAT	1530	
Db	6399	CATATGATACGAGAGTACATATGTATGTATGTATGTATGTATGTATGTATGTAT	6458	
Qy	1531	ACCCACAAGAAAGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTAT	1590	
Db	6459	ACCCACAAGAAAGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTAT	6518	
Qy	1591	TGATAGACAGATGATGAGATATATATATATATATATATATATATATATATATATAT	1650	
Db	6519	TGATAGACAGATGATGAGATATATATATATATATATATATATATATATATATATAT	6578	
Qy	1651	TAAATTTAACCCACTCTGTGTATGTATGTATGTATGTATGTATGTATGTATGTAT	1710	
Db	6579	TAAATTTAACCCACTCTGTGTATGTATGTATGTATGTATGTATGTATGTATGTAT	6638	
Qy	1711	CCATATAGATAGAGGGGAGATATATATATATATATATATATATATATATATATAT	1770	
Db	6639	CCATATAGATAGAGGGGAGATATATATATATATATATATATATATATATATATAT	6698	
Qy	1771	ATATACAGACAGATGAAGATATAGGTGACAGAAAGATATATATATATATATATAT	1830	
Db	6699	ATATACAGACAGATGAAGATATAGGTGACAGAAAGATATATATATATATATATAT	6758	
Qy	1831	AT	1890	
Db	6759	AT	6818	
Qy	1891	CACAGGCTGTCCAAAGGTATCCCTTGAAGCAATTCATATATATATATATATATAT	1950	
Db	6819	CACAGGCTGTCCAAAGGTATCCCTTGAAGCAATTCATATATATATATATATATAT	6878	
Qy	1951	GTTTTGGCATCTTAAAT	2010	
Db	6879	GTTTTGGCATCTTAAAT	6938	
Qy	2011	TCACACAGATACAT	2070	
Db	6939	TCACACAGATACAT	6998	
Qy	2071	ATGGCAGTCTGAGAGAGAGAGATATATATATATATATATATATATATATATATAT	2130	
Db	6999	ATGGCAGTCTGAGAGAGAGAGATATATATATATATATATATATATATATATATAT	7058	
Qy	2131	AAACCAT	2190	
Db	7059	AAACCAT	7118	
Qy	2191	ATACAG	2250	
Db	7119	ATACAG	7178	
Qy	2251	AAATAG	2310	
Db	7179	AAATAG	7238	
Qy	2311	TAAATAG	2370	
Db	7239	TAAATAG	7298	

Qy	2371	AGCAATCCTCAGAGAGGGGACCCAGAAATTTGTAACGCACTTTTAAATTTGAGGGGAAT	2430	
Db	7299	AGCAATCCTCAGAGAGGGGACCCAGAAATTTGTAACGCACTTTTAAATTTGAGGGGAAT	7358	
Qy	2431	TTTTTCTACTGTAATTTCAACACACTGTTTAAATAGTACTTGTTTAAATAGTACTTG	2490	
Db	7359	TTTTTCTACTGTAATTTCAACACACTGTTTAAATAGTACTTGTTTAAATAGTACTTG	7418	
Qy	2491	CTGAAGGCTCAAAATACAGTGAAGAGAGTGCACAAATCACACCTCCATGCAATAAAC	2550	
Db	7419	CTGAAGGCTCAAAATACAGTGAAGAGAGTGCACAAATCACACCTCCATGCAATAAAC	7478	
Qy	2551	AAATTTAAACATGTGGCAGAGAAATAGAAAAGCAATGTATGCCCTCCATCAGTGGAC	2610	
Db	7479	AAATTTAAACATGTGGCAGAGAAATAGAAAAGCAATGTATGCCCTCCATCAGTGGAC	7538	
Qy	2611	AAATTAATGTCTCAATATTTACTGGCTGCTATTAACAAGAGATGTGTAAATACA	2670	
Db	7539	AAATTAATGTCTCAATATTTACTGGCTGCTATTAACAAGAGATGTGTAAATACA	7598	
Qy	2671	ACAATGGCTCCAGATCTTCAAGACCTGAGAGAGCGATATAGAGGACAATTTGGAAGTG	2730	
Db	7599	ACAATGGCTCCAGATCTTCAAGACCTGAGAGAGCGATATAGAGGACAATTTGGAAGTG	7658	
Qy	2731	AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2790	
Db	7659	AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	7718	
Qy	2791	AGAAAGAGTGTGCACAGAGAAAGAAAGACAGATGGGAATAGAGGCTTTGCTTGGGT	2850	
Db	7719	AGAAAGAGTGTGCACAGAGAAAGAAAGACAGATGGGAATAGAGGCTTTGCTTGGGT	7778	
Qy	2851	TCTTGGAGCAGCAGAGAGACATATGGGCTGCACGTCAATGACCTGACGGTACAGGCCA	2910	
Db	7779	TCTTGGAGCAGCAGAGAGACATATGGGCTGCACGTCAATGACCTGACGGTACAGGCCA	7838	
Qy	2911	GACATTTATTTGTCTGATATATAGTGCACAGAGAAATTTTGTCTAGGGCTATTGAGGCG	2970	
Db	7839	GACATTTATTTGTCTGATATATAGTGCACAGAGAAATTTTGTCTAGGGCTATTGAGGCG	7898	
Qy	2971	AACGACATCTGTTCAACCTCAGACTGAGGCAATCAACAGCTCCAGGCAAGAAATCTCTG	3030	
Db	7899	AACGACATCTGTTCAACCTCAGACTGAGGCAATCAACAGCTCCAGGCAAGAAATCTCTG	7958	
Qy	3031	CTGTGAGAAAGATACCTAAAGGATCAACAGCTCTGGGGATTTGGGGTGTCTTGGAATC	3090	
Db	7959	CTGTGAGAAAGATACCTAAAGGATCAACAGCTCTGGGGATTTGGGGTGTCTTGGAATC	8018	
Qy	3091	TCATTTGCACACCTGCTGTGCTTGGAAATGCTAGTTGGAGTAAATTTCTTGTGAACAGA	3150	
Db	8019	TCATTTGCACACCTGCTGTGCTTGGAAATGCTAGTTGGAGTAAATTTCTTGTGAACAGA	8078	
Qy	3151	TTTGGAAATACATACCTGGAATGGAGTGGGACAGAGAAATTTAAATTAACAAGCTTAA	3210	
Db	8079	TTTGGAAATACATACCTGGAATGGAGTGGGACAGAGAAATTTAAATTAACAAGCTTAA	8138	
Qy	3211	TACACTCCTTAATTAAGAAATCGCAAAACAGCAAGAAAGAAATGAACAAGAAATTTATG	3270	
Db	8139	TACACTCCTTAATTAAGAAATCGCAAAACAGCAAGAAAGAAATGAACAAGAAATTTATG	8198	
Qy	3271	AAATAGATTAATTTGGGCAAGTTTGTGAATTTGTTTAAACATPAACAATTTGCTGTATTA	3330	
Db	8199	AAATAGATTAATTTGGGCAAGTTTGTGAATTTGTTTAAACATPAACAATTTGCTGTATTA	8258	
Qy	3331	TAAATTTATTCATTAATGATAGTACAGAGCTTTGGATTTAAAGAAATAGTTTCTCTATAC	3390	
Db	8259	TAAATTTATTCATTAATGATAGTACAGAGCTTTGGATTTAAAGAAATAGTTTCTCTATAC	8318	
Qy	3391	TTTCTATAGTAAATAGATTAAGCAGGAGATTTACCATTTATGCTTTACAGCCACCTCC	3450	
Db	8319	TTTCTATAGTAAATAGATTAAGCAGGAGATTTACCATTTATGCTTTACAGCCACCTCC	8378	

```

Oy 3451 CAATCCGAGGGGACCCGACAGGCGCCGAAGGATAGAAGAGAGTGGAGAGACA 3510
Db 8379 CAATCCGAGGGGACCCGACAGGCGCCGAAGGATAGAAGAGAGTGGAGAGACA 8438
Oy 3511 GAGACAGATCCATTGATAGTGAACGGATCCTTAGCATTTATCTGGGACGATCTGGCGA 3570
Db 8439 GAGACAGATCCATTGATAGTGAACGGATCCTTAGCATTTATCTGGGACGATCTGGCGA 8498
Oy 3571 GCGTGTGCTCTTACGCTACCAAGCGCTTGAAGAGCTTACTGCTTAATGCTTAAGAGATGG 3630
Db 8499 GCGTGTGCTCTTACGCTACCAAGCGCTTGAAGAGCTTACTGCTTAATGCTTAAGAGATGG 8558
Oy 3631 TGAAGCTTCTGGAGCGAGGGGGTGGGAAGCCCTCAATATGTGGGAATGCTCTACAGT 3690
Db 8559 TGAAGCTTCTGGAGCGAGGGGGTGGGAAGCCCTCAATATGTGGGAATGCTCTACAGT 8618
Oy 3691 ATTGAGTACAGAACTAAGAAATAGTGTCTTAACTTCTCAATGCGACAGCCATATAGCAG 3750
Db 8619 ATTGAGTACAGAACTAAGAAATAGTGTCTTAACTTCTCAATGCGACAGCCATATAGCAG 8678
Oy 3751 TAGCTGAGGGGACAGATAGGGTTATAGAGTATTACAAGCAGCTTATAGAGCTATTCGCC 3810
Db 8679 TAGCTGAGGGGACAGATAGGGTTATAGAGTATTACAAGCAGCTTATAGAGCTATTCGCC 8738
Oy 3811 ACATACCTAGAAAGATAGAAGACAGGCGCTTGAAGAGATTTGCTATAGAGTGGTGGCAAG 3870
Db 8739 ACATACCTAGAAAGATAGAAGACAGGCGCTTGAAGAGATTTGCTATAGAGTGGTGGCAAG 8798
Oy 3871 TGTGTCAAAAAGTAGTGTGATGTGATGCTCTGCTTAAGGAGAAAGATAGAGCAGCTGAG 3930
Db 8799 TGTGTCAAAAAGTAGTGTGATGTGATGCTCTGCTTAAGGAGAAAGATAGAGCAGCTGAG 8858
Oy 3931 CCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAGACTAGA 3971
Db 8859 CCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAGACTAGA 8899

RESULT 5
LOCUS HIVNL43 9709 bp ss-RNA linear VRL 02-AUG-1993
DEFINITION Human immunodeficiency virus type 1, NY5/BRU (LAV-1) recombinant
clone pNL4-3.
ACCESSION M19921
VERSION M19921.1 GI:328415
KEYWORDS Human immunodeficiency virus 1 (HIV-1)
SOURCE Human immunodeficiency virus 1
ORGANISM
    Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
    lentivirus group.
REFERENCE 1 (bases 1 to 9709)
AUTHORS Adachi,A., Gendelman,H.E., Koenig,S., Folks,T., Willey,R.,
TITLE Radson,A. and Martin,M.A.
JOURNAL Production of acquired immunodeficiency syndrome-associated
MEDLINE retrovirus in human and nonhuman cells transfected with an
PUBMED infectious molecular clone
JOURNAL J. Virol. 59 (2), 284-291 (1986)
REFERENCE 2 (bases 1 to 9709)
AUTHORS Buckler,C.E., Buckler-White,A.J., Willey,R.L. and McCoy,J.
TITLE Unpublished (1988)
JOURNAL 3 (sites)
AUTHORS Dai,L.C., West,K., Littau,R., Takahashi,K. and Ennis,F.A.
TITLE Mutation of human immunodeficiency virus type 1 at amino acid 585
on gp1 results in loss of killing by CD8+ A24-restricted cytotoxic
T lymphocytes
JOURNAL J. Virol. 66 (5), 3151-3154 (1992)
MEDLINE 92219406
PUBMED 1373204
COMMENT Original source text: Human immunodeficiency virus type 1 (HIV-1),
NY5/BRU (LAV-1) recombinant clone pNL4-3.
[3] sites; revisions of [3].
Clean copy of sequence [3] kindly provided by Chuck Buckler, NIAID,

```

FEATURES

source

Bethesda, MD, 24-JUN-1988. The construction of pNL4-3 has been described in [1]. pNL4-3 is a recombinant (infectious) proviral clone that contains DNA from HIV isolates NY5 (5' half) and BRU (3' half). The site of recombination is the EcoRI site at positions 5743-5748. The length and sequence of the vpr coding region corresponds to that of the BRU, SC, SP2, MAL, and ELI isolates. The vpr coding region of these isolates is about 18 amino acid residues longer than the vpr coding region of the IIB isolates. In HIVNL43, this shift is due to a single base deletion (with respect to the IIB's) at position 5770. The sequence at this position is 'attc' in HIVNL43 and 'atttc' in HIVXB2. The original BRU clone, sequenced by Martin-Hobson, et al. (Cell 40, 9-17 (1985)), and the BRU portion of the pNL4-3 recombinant clone are different clones from the same BRU isolate. Two of the revisions reported in the FEATURES produced changes in amino acid sequences. The revision at position 2421 changes one amino acid residue from 'R' to 'G' in the pol coding region. The revision at positions 8995-9000 changes three amino acid residues from 'AHT' to 'VTP' in the nef coding region.

Location/Qualifiers

1..9709

/organism="Human immunodeficiency virus 1"

/mol_type="genomic RNA"

/db_xref="taxon:11676"

1..634

/note="5', LTR"

454..550

/note="R repeat 5' copy"

455..9626

/note="tat, rev, nef subgenomic mRNA"

744..5776

/note="tat, rev, nef mRNA intron 1"

790..2292

/note="gag polyprotein"

/codon_start=1

/protein_id="AA44987.1"

/db_xref="GI:328418"

/translation="WGARASVLSGSELDKMKELIRPGCGKQYLYLKIIVASRELEF

AVNPGLETSBGRQILGOLPDSLSSELSLYNTIAYLCVHORIDVDRLEAD

KIEEONKSKKKAQAADADNGNSQVNPYVONLOGOMHQAISPTINAWKLYVE

EKASPEVIMPEFSALSEGAPDQNTMLNTVGHQAMOMLKEITINEEAEMDLHYE

HAGPIARGOMREPRSDIAGTSTLQDGMTPNPILPGEITKYLIIINGLKIVM

YSPSTIIDIRGQKPREPRDYDRPYRTIARAQASQYKMMKMTETVIAKNNITGK

KALGEGATLEBMTACOGVGPGRKAVYLAEMASQVTPATIMIQKNFRQRTVC

FNCGRKEHIAKNCRAPRKGCWCRCGKHQMKDCTERQANFLGIMPSHKGPRGNFLQ

SRPEPTAPPESEFRFGETTTPSQKQEPIDKELYPLASLSLFGSDPSQ"

2085..5096

/partial

/note="pol polyprotein (NH2-terminus uncertain)"

/codon_start=1

/protein_id="AA44988.1"

/db_xref="GI:328419"

/translation="FFREDLAIPQKAREFSEEQTRANSPTRELQVAGRDNSLSEA

GADRGQVFSFSDIITLMORPLVITKIGGOLKEALDPTGADDTLEEMNLPGMKPM

IGCGIFRIKYNQDOLILEICGHAKICTVAGPVPVNIIGRLITQICITLPIPSI

ETPVYKLRPGADGPKVAKOMPLETEKIALVETCTEMKESGKISTGEPNPTPVRI

KKQSTWRKLVDFRELNRKTRDQWEVQLGIPRAGLQKQSVTVLDVGAVESVPLD

KDFRYTAFTPIINNETPGIRYOYENVLPQGWKSPAIIFQCSMTKILEPRKQNDIV

IYOYMDLYQSDLEIGONHRTKLELQNLHLMFTTPDKKHQEPFLMGMYLELHD

KMTVPQVILPEKDSMTVONDKLQKLVNMAISOYIAGIKVROLCKLGTALTEVPL

TEBAELFLAKRELITKPRVHCYVYDPSKDLAETQKQSGCOMPTQIYOEPKNTKTK

YARKGKHTNDVQKQLEAVOKIATESYIVMGKTPKELPTQKETEMAWMEIYQATVI

PEWEVTPPLVYKIMYOLEKEPIIGATFVYVDGANETKIGKGYVTDGRQKVVPL

TDITNOKTELOAHILADSGLEVNITVDSQVALGIIQADPKSESLSVLIQLIK

KEKYVLAMPVANKHIGNEQVDGLVSGIRKVLFDLQIDKAEHEKVLHSMRAMASQ

FNLPVPAKEIVASCDCOLKGEAMHGOVQSPQIMQDCTHLEGKYLIVAVHVASQY

IEAEVPAETGOETAPFLKLAGEMPVKTVHTQNSMFTSTTVAAACMGVAKIOEFT

PTNFQSGVIESNKKELKITIGYRQDAENHKTRVQMAVPIHNFRKRGCGIGYSAGER

IVDLIADIQTELOKQITKQIRFVYVIRSDRQVDMGPAKLLKMGGAVALVQNSDI

KVPRRRAKIIRDKQMGADDDCVASRDD"

5041..5619

/note="vif protein"

CDS

[illegible]

LTR		976.. .9709 /note="3' LTR"	
repeat_region		9529.. .9626 /note="R repeat 3' copy"	
polyA_signal		9602.. .9607 /note="mRNA polyadenylation signal"	
BASE COUNT	3421 a	1756 c	2366 g 2166 t
ORIGIN	5'	terminus of NY5 LTR.	
Query Match		43.0%; Score 2676.2; DB 14;	Length 9709; Best Local Similarity 99.9% Matches 2678:
		Conservative 0;	Mismatches 3; Indels 0; Gaps 0;
Dy	1291	CCATGAGTGAAGAAGTAGATACAGCACTTTGTGGAGATGGGGTGAAATGGGCACCA	1350
Dd	6219	CAATTAGAGTGAAGAGACAATCAGACTTCGTGGAGATGGGGGTGAAAATGGGGCACCA	6278
OY	1351	TGCTCCTTGGGATNTGATCTGTCTGTAAGATAAGAAAAATPTGGGTACCCTCTATT	1410
Dd	6279	TGCTCCTTGGGATNTGATCTGTCTGTAAGATAAGAAAAATPTGGGTACCCTCTATT	6338
OY	1411	ATGGGGTACCTGTGTGGAAGAACACACCACTCATTTTTGTGATCAGATGCTTAAG	1470
Dd	6339	ATGGGGTACCTGTGTGGAAGAACACCACTCATTTTTGTGATCAGATGCTTAAG	6398
OY	1471	CATATGATACAGGGTACATATGTTTGGGCCACACATGCTGTGTACCCACAGACCCA	1530
Dd	6399	CATATGATACAGGGTACATATGTTTGGGCCACACATGCTGTGTACCCACAGACCCA	6458
OY	1531	AOCACAAAGATGTGTATGTGTAANTGACAGAAAAATTTAACATGTGGAAAAATACA	1590
Dd	6459	AOCACAAAGATGTGTATGTGTAANTGACAGAAAAATTTAACATGTGGAAAAATACA	6518
OY	1591	TGATGAGACAGATCGATGAGATATTAATCAGTTTATGGATCAAAGCCTTAAGCATGTG	1650
Dd	6519	TGATGAGACAGATCGATGAGATATTAATCAGTTTATGGATCAAAGCCTTAAGCATGTG	6578
OY	1651	TAAATTTAACCCCACCTCTGTGTAGTTTAAAGTGCACATGATTTGAGAATGATCTATA	1710
Dd	6579	TAAATTTAACCCCACCTCTGTGTAGTTTAAAGTGCACATGATTTGAGAATGATCTATA	6638
OY	1711	CCATAGTAGTAGCGGGAGAAATGATATGAGAAAGAGAGATAAAAAAGTCTCTTTCA	1770
Dd	6639	CCATAGTAGTAGCGGGAGAAATGATATGAGAAAGAGAGATAAAAAAGTCTCTTTCA	6698
OY	1771	ATATCAGACAGCATTAGAGATTAAGGTGCGAAGAAATATGCAATTCCTTTATTAACCTG	1830
Dd	6699	ATATCAGACAGCATTAGAGATTAAGGTGCGAAGAAATATGCAATTCCTTTATTAACCTG	6758
OY	1831	ATATAGTACCAATAGATTAATACCAAGCTATAGGTTGATPAAGTTGTACACCTCACTCATTA	1890
Dd	6759	ATATAGTACCAATAGATTAATACCAAGCTATAGGTTGATPAAGTTGTACACCTCACTCATTA	6818
OY	1891	CACAGGCTGTGCCAAAGSTATCTTTTGAGCCAAITCCCATTCATTTATTTGCCCCGGCTG	1950
Dd	6819	CACAGGCTGTGCCAAAGSTATCTTTTGAGCCAAITCCCATTCATTTATTTGCCCCGGCTG	6878
OY	1951	GTTTTGCATTTCTAAATGTATPATATAGAGCTTCAATGGAAGGAGCCATGTCAATG	2010
Dd	6879	GTTTTGCATTTCTAAATGTATPATATAGAGCTTCAATGGAAGGAGCCATGTCAATG	6938
OY	2011	TCAGACAGTCAATGTACACATGGAATCAGGCCAGTAGATTCAACTCAACTGCTGTAA	2070
Dd	6939	TCAGACAGTCAATGTACACATGGAATCAGGCCAGTAGATTCAACTCAACTGCTGTAA	6998
OY	2071	ATGCGAGTCTAGCAGAGAAGATGTAGTAATTTAGATCTGCCAATTTTCACAGACATGCTA	2130
Dd	6999	ATGCGAGTCTAGCAGAGAAGATGTAGTAATTTAGATCTGCCAATTTTCACAGACATGCTA	7058
OY	2131	AAACCATATAGTACAGCTGAACAACATCTGTGANAATTAATTTGNACAAGACCCACAACA	2190
Dd	7059	AAACCATATAGTACAGCTGAACAACATCTGTGANAATTAATTTGNACAAGACCCACAACA	7118

OY 2191 ATCAAGAAAAAGTATCCGTATCCAGAGGGAGCCAGGAGACATTTGTACAAATAGGAA 2250
 |||||||
 Db 7119 ATACAGAAAAAGTATCCGTATCCAGAGGGAGCCAGGAGAGCATTTGTACAAATAGGAA 7178
 OY 2251 AAATAGGAATATGAGACAGACATTTGTAACATTTAGTAGAGCAAAATGGAATGCCACTT 2310
 |||||||
 Db 7179 AAATAGGAATATGAGACAGACATTTGTAACATTTAGTAGAGCAAAATGGAATGCCACTT 7238
 OY 2311 TAAAGAGATAGCTAGCAAAATTTAGAGAGCAAAATTTGGAAATATATTAATTAATCTTTA 2370
 |||||||
 Db 7239 TAAAGAGATAGCTAGCAAAATTTAGAGAGCAAAATTTGGAAATATATTAATTAATCTTTA 7298
 OY 2371 AGCAATCCTCAGAGAGGGAGCCAGAAATTTGACGACAGTTTTAATTTGTGAGGGAGAT 2430
 |||||||
 Db 7299 AGCAATCCTCAGAGAGGGAGCCAGAAATTTGACGACAGTTTTAATTTGTGAGGGAGAT 7358
 OY 2431 TTTTCTACTGTAAATTCACACACACTGTTTAAATAGTACTTGGTTTAAATAGTACTTGAAGTA 2490
 |||||||
 Db 7359 TTTTCTACTGTAAATTCACACACACTGTTTAAATAGTACTTGGTTTAAATAGTACTTGAAGTA 7418
 OY 2491 CTGAAGGTCATAATTAACCTGAAGAGAGTGCACACATTCACCTCCCATGCAAGATTAAC 2550
 |||||||
 Db 7419 CTGAAGGTCATAATTAACCTGAAGAGAGTGCACACATTCACCTCCCATGCAAGATTAAC 7478
 OY 2551 AATTATTAACATGTGGCAGAGAGTGAAGAAAGCAATGTATGCCCTCCCATGCAAGTGAAC 2610
 |||||||
 Db 7479 AATTATTAACATGTGGCAGAGAGTGAAGAAAGCAATGTATGCCCTCCCATGCAAGTGAAC 7538
 OY 2611 AAATTAAGTGTTCATCAAAATATTAAGTGGCTGTATTAAACAAGAGATGGTGAATTAACA 2670
 |||||||
 Db 7539 AAATTAAGTGTTCATCAAAATATTAAGTGGCTGTATTAAACAAGAGATGGTGAATTAACA 7598
 OY 2671 ACAATGGTCCGAGATCTTCAGACCTGGAGAGGCGATATGAGGAGCAATGGGAGAGTGTG 2730
 |||||||
 Db 7599 ACAATGGTCCGAGATCTTCAGACCTGGAGAGGCGATATGAGGAGCAATGGGAGAGTGTG 7658
 OY 2731 AATTATTAATTAATTAAGTACTTAAATTAAGTGAACCATTAAGAGTGAACCCCAAGGCAA 2790
 |||||||
 Db 7659 AATTATTAATTAATTAAGTACTTAAATTAAGTGAACCATTAAGAGTGAACCCCAAGGCAA 7718
 OY 2791 AGAAGAGAGTGTGCAG 2850
 |||||||
 Db 7719 AGAAGAGAGTGTGCAG 7778
 OY 2851 TCTTGGGAGCAG 2910
 |||||||
 Db 7779 TCTTGGGAGCAG 7838
 OY 2911 GACAATTAATTAATTAAGTACTTAAATTAAGTGAACCATTAAGAGTGAACCCCAAGGCAA 2970
 |||||||
 Db 7839 GACAATTAATTAATTAAGTACTTAAATTAAGTGAACCATTAAGAGTGAACCCCAAGGCAA 7898
 OY 2971 AACAGCATCTTGTGCACTCAGAGTGTGGGCAATCAACAGCTCCAGGCAAGAAATCTCTG 3030
 |||||||
 Db 7899 AACAGCATCTTGTGCACTCAGAGTGTGGGCAATCAACAGCTCCAGGCAAGAAATCTCTG 7958
 OY 3031 CTGAGAGAAAGATTAAGATCAACAGCTCTGGGAGATTTGGGGTGTCTGGAAGAAC 3090
 |||||||
 Db 7959 CTGAGAGAAAGATTAAGATCAACAGCTCTGGGAGATTTGGGGTGTCTGGAAGAAC 8018
 OY 3091 TCATTTGACACCATCTGTGCTTGTGATGCTAGTGTGAGTAATTAATCTCTGGAAGAAC 3150
 |||||||
 Db 8019 TCATTTGACACCATCTGTGCTTGTGATGCTAGTGTGAGTAATTAATCTCTGGAAGAAC 8078
 OY 3151 TTTTGAATTAATTAAGATCAACAGCTCTGGGCAATCAACAGCTCCAGGCAAGAAATCTCTG 3210
 |||||||
 Db 8079 TTTTGAATTAATTAAGATCAACAGCTCTGGGCAATCAACAGCTCCAGGCAAGAAATCTCTG 8138
 OY 3211 TACACTCTTAAATTAAGATCAACAGCTCTGGGCAATCAACAGCTCCAGGCAAGAAATCTCTG 3270
 |||||||
 Db 8139 TACACTCTTAAATTAAGATCAACAGCTCTGGGCAATCAACAGCTCCAGGCAAGAAATCTCTG 8198
 OY 3271 AATTAGATTAATTAAGATCAACAGCTCTGGGCAATCAACAGCTCCAGGCAAGAAATCTCTG 3330

Db 8199 AATTAGATTAATTAAGATCAACAGCTCTGGGCAATCAACAGCTCCAGGCAAGAAATCTCTG 8258
 |||||||
 OY 3331 TAAATTAATTAATTAAGATCAACAGCTCTGGGCAATCAACAGCTCCAGGCAAGAAATCTCTG 3390
 |||||||
 Db 8259 TAAATTAATTAATTAAGATCAACAGCTCTGGGCAATCAACAGCTCCAGGCAAGAAATCTCTG 8318
 OY 3391 TTTTCTATTAATTAAGATCAACAGCTCTGGGCAATCAACAGCTCCAGGCAAGAAATCTCTG 3450
 |||||||
 Db 8319 TTTTCTATTAATTAAGATCAACAGCTCTGGGCAATCAACAGCTCCAGGCAAGAAATCTCTG 8378
 OY 3451 CAATCCGAGAGGAGCCGACAGCCGCAAGAGATTAAGAGAGAGTGTGAGAGAGACA 3510
 |||||||
 Db 8379 CAATCCGAGAGGAGCCGACAGCCGCAAGAGATTAAGAGAGAGTGTGAGAGAGACA 8438
 OY 3511 GAGACAGATTCATTCGATTAAGAGAGATTCGATTAAGAGAGATTCGATTAAGAGAGATTCG 3570
 |||||||
 Db 8439 GAGACAGATTCATTCGATTAAGAGAGATTCGATTAAGAGAGATTCGATTAAGAGAGATTCG 8498
 OY 3571 GCTGTGCTCTTACGATTAAGAGAGATTCGATTAAGAGAGATTCGATTAAGAGAGATTCG 3630
 |||||||
 Db 8499 GCTGTGCTCTTACGATTAAGAGAGATTCGATTAAGAGAGATTCGATTAAGAGAGATTCG 8558
 OY 3631 TGAACCTCTGAG 3690
 |||||||
 Db 8559 TGAACCTCTGAG 8618
 OY 3691 ATGGAGTCAAG 3750
 |||||||
 Db 8619 ATGGAGTCAAG 8678
 OY 3751 TAGCTGAG 3810
 |||||||
 Db 8679 TAGCTGAG 8778
 OY 3811 ACATACCTAG 3870
 |||||||
 Db 8739 ACATACCTAG 8798
 OY 3871 TGGTCAAAAAGTATGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3930
 |||||||
 Db 8799 TGGTCAAAAAGTATGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 8858
 OY 3931 CCAGCAGCAGATGGGAG 3991
 |||||||
 Db 8859 CCAGCAGCAGATGGGAG 8999

RESULT 6
 AF324493
 LOCUS 14824 bp DNA circular SYN 15-FEB-2001
 DEFINITION HIV-1 vector pNL4-3, complete sequence.
 ACCESSION AF324493
 VERSION AF324493.1 GI:12831134
 KEYWORDS
 SOURCE
 ORGANISM
 HIV-1 vector pNL4-3
 HIV-1 vector pNL4-3
 artificial sequences; vectors.
 REFERENCE
 1 (bases 1 to 9709)
 Adachi,A., Gendelman,H.E., Koenig,S., Folks,T., Willey,R.,
 Rabson,A. and Martin,M.A.
 Production of acquired immunodeficiency syndrome-associated
 retrovirus in human and nonhuman cells transfected with an
 infectious molecular clone
 J. Virol. 59 (2), 284-291 (1986)
 MEDLINE
 86281827
 PUBMED
 3016298
 2 (bases 1 to 14824)
 Bosche,W.J., Poon,D.T.K., Ott,D.E., Hu,W.-S. and Gorelick,R.J.
 Complete Plasmid Sequence of pNL4-3
 JOURNAL
 Unpublished
 3 (bases 1 to 14824)
 Bosche,W.J., Poon,D.T.K., Ott,D.E., Hu,W.-S. and Gorelick,R.J.

TITLE	
JOURNAL	Direct Submission
Submitted (28-NOV-2000) AIDS Vaccine Program, SAIC Frederick and Drug Resistance Program, National Cancer Institute at Frederick, PO Box B, Frederick, Maryland 21702-1202, USA	
FEATURES	
source	Location/Qualifiers
	1. .14824
	/organism="HIV-1 vector pNL4-3"
	/mol_type="other DNA"
	/db_xref="taxon:151458"
	/plasmid="pNL4-3"
	/focus
source	1. .9709
	/organism="Human immunodeficiency virus 1"
	/proviral
	/mol_type="other DNA"
	/strain="NY5/BRU (LAV-1)"
	/db_xref="taxon:11676"
	/note="from sequence deposited in GenBank Accession Number M19921"
LTR	1. .634
	/note="5' long terminal repeat"
	454. .550
repeat_region	/note="R: 5' copy"
	790. .2292
CDS	/codon_start=1
	/product="gag polyprotein"
	/protein_id="AAK08483.1"
	/db_xref="GI:12831136"
	/translation="MGARASVLSGGELDKMKITLRPGGKKQYKLIKHWASRLERF AVNPGILBETSEGRQILIGQLPQLQTSSEELRSYNTIAVLYCVHQRIADVKTLEAD KIEEONKSKRKAQAADGNNNSQVSNPVIYONLOGOVHOAISPTLANAWKVE EKASPEVIMPFMSALSEGATPODLNMTLNTVGHOAMOMLKETINBEAEWDLHVE HAGIAGCOMREPRGSDIAGTTSTLOIGIMTNPINPPIEGEIKRWIILGLNKIVRY XSPSTIIDRGGPEPRDYDRFKTLRAFOASOEKNNMTETLLONKFNROCKTTL KALGPGATLEEMTNACOGVCGPGKARVLAFAOASQVYNPATIMLOKFNROCKTVC PNCKEGHIAKNCRAPKRGCKGCKGSHOKDKTERQANFLGKIWPSHGKRPENFLQ SRPEPTAPPESEFRFGEETTPSQOEPRIDELYPPLASLRSLSFGSDPSQ"
	<2085. .5096
	/codon_start=1
	/transl_table=1
	/product="pol polyprotein"
	/protein_id="AAK08484.1"
	/db_xref="GI:12831137"
	/translation="FFREDLAFPOGKAREFSSSEQTRANSPTRRELQVGRDNLSBA GADRGVSEFSPQITLMORPLVTIKIGGOLKEALDLPDADTVLEEMNLVGRKPAK IGGIGFKVROVYQIILITCGHKAIGTVLGPVNIIGNILITLQIGCTINPISPI ETVPYVKLRPGMDGPKVQMPLETEKIKALVPTCTMEKECKISKIGPBNPTNTPVRI KKDSTIKWRLVDRRELNKRTODEWEVOLGIPHPAGLKOKKSVTLVDGVDFSVPLD KDFRKYAFITPSINNETPGIRYOYNYLPQMGKSPAIFFQCSMTKILEPFRKQNDIV IYQYMDLYVSDLEIGQHRKIEBLRQHLRMGFTTPDKKHOKPEPFLMMGYELHPD KWTQPIVLPEKDSMTVNDIQKLKVLKLMASQIYAGIKRQLCKLRTGKALTYVPL TEEAELEAENREILKEPVGVYVDPKSDLAETIOKOGOGOMTQIYOEPKRLKTVK YARKGAHTNDVROLTEAVOQILATESIYVINGKTPRKLPILOKETWEAMWTEYNATNI PEREFVNTPPVYKLTQLEKEPTITGAETFYDGAENETKLGKAGIYTDROKRVPL TDTTNOKTELOAHIALQDSGLEVINIYDSOYALGIIQADPKSESELVSOILEQLK KEKYLVAVPAHAKGIEGNEVDVDSAGIRVYFLDIDKAQEHKEHYHSNWRAMAD FNLPPVVAKEIVASCDCQOLGEMHGQVDCPSIQMDLCTHLEGVILVAVASQY IEAEVIPAETGSOETAYFLKLAGRPVKTYTONGSNFTSTTVCAACMGVAKOEFGI PYNOSOGVIESMEKELKTIIGVRODAENHKTAVOMAVFIHNRKRGKGGYAGSAGER IVDIATDIOETKELOKITKIQNFRVYTRPSRDPVWKGPAKLKMGSGAVYTQNSDI KVVPRRAKIIIRDYKQMAQDDCVASRODEB"
	2421
misc_difference	/note="Compared to sequence deposited in GenBank Accession Number M19921: results in arginine to glycine substitution"
	/replacement="g"
	5041. .5619
CDS	/codon_start=1
	/transl_table=1
	/product="yif protein"
	/protein_id="AAK08482.1"
	/db_xref="GI:12831135"

CDS	/translation="MENRMQVMIWVOYDRNRINTWKLKYNHMYISRKAKDWPYRHNY ESNPKLSSSEVHPIPLGDAKLVTITWYMLHTEGDMHGGQVSIEMRKRYSTQVDDL ADOLIHLYHDCSESAIRNTILGRIYSPCEYQACHNKGSLQYLALAILIKPKQIK PPLPSYVKLETFEDRMNKPQTKRGHSGTNGH"
	5559. .5849
	/codon_start=1
	/transl_table=1
	/product="vpr protein"
	/protein_id="AAK08485.1"
	/db_xref="GI:12831138"
	/translation="MEQAPPDQGPQRPYMEWTLLEELKESEAVHFRPIWTLNLQO HITTEGTDWAGDAITRILOQLTFHFRIGCHRSRIGTVRORARNGASRS"
	5743. .5748
misc_feature	/note="EcoRI site of recombination"
	join(5830. .6044,8369. .8414)
CDS	/codon_start=1
	/transl_table=1
	/product="tat protein"
	/protein_id="AAK08486.1"
	/db_xref="GI:12831139"
	/translation="MEPVDRLEPMKHHGSOPTACTNCYCKCFHCQYCFMTKALG ISYGRKRRORRRARHONSORTHOASLSQPTSQSRGDPGPK"
	join(5969. .6044,8369. .8643)
CDS	/codon_start=1
	/transl_table=1
	/product="rev protein"
	/protein_id="AAK08487.1"
	/db_xref="GI:12831140"
	/translation="MAGRSQSDSEELRTVRLIKLKYOSNPPNPESTRQARRRRR WRERODHISERILSTYLGRSHEPVLPDLPRLRLTDCNECGTSGTGQVGSPOI LVESPTGEGTE"
	6061. .6306
CDS	/codon_start=1
	/transl_table=1
	/product="vpu protein"
	/protein_id="AAK08488.1"
	/db_xref="GI:12831141"
	/translation="MOPRIYAIYALVVAIIIAIVWSIVLIEYRKILIKQRKIDRLDR LIERAEDSGNESEGVSAIVEMGVEMGHAPWIDDL"
	6221. .8785
	/codon_start=1
	/transl_table=1
	/product="envelope polyprotein"
	/protein_id="AAK08489.1"
	/db_xref="GI:12831142"
	/translation="MYRKEKYOHLMRMGKMGTMGLILMGCSATEKLWTVYGVYV WKATYTIKASQAKADPDEVHVMYATHACVPTDPNQEVLVNTNEMKENDVDE OMHEDTISLMDQSLKPCVKITPLCYSLKCTDLKNDNTNNSGSMIMEKECIKCSFN ISTSIRKVGKERAFFYKLDIVPIDNTSYRLISCNSTVITQACPKVSFEPIPIHYCAP AGAAILKCNKRTGTCPTNVSIVQCTHGIQRYVSTQLLNGSLAEEDVIRISANT DNKATILVQNTSVEINCTPNNNTRKSIQIRGPGAFYTIKIGMROACNISRA KMNATLQWISKILREOQGNKTTIFKQSSGDEIYHSFNGGEFFYCNSTOLFNSY WPNSTQSTBESNNGSDTTLTCRIKQPIINMOEGCKAMVADPISQIJCSSNTLGL LTRDGGNNNNGSEIIRPGGDMRDNRSELTYKYKTEPLDVAIPKARRVYOGEK RAYGALFLGFLGAGSTMGCTSMSTLTVAORULSDIVOOQNNLNALEAOHLLQD TVGICKOLQRIILAVERYLDQDLGLTGWGSGKLICTTAVPMAWSMSKSLDEIMNM TMWMEDEINNYTSLISLSEENQOGEKDEBELLEDKAASLMNFNTIMMYLITKL FIMIVGSLVCLRIYFAVLSIYNRQOQSPSLSTQTHPIRGDRPREGIEEGEGEDR DRSIRLVNGSLALIMDLRSICLFSYHRLDLILLIYRIYELHGRGWEKLYKMNML QYSGOELKNSAVNLLNATLAVAEGRDRIYEVLDAAVRAIRHPIRRIQGLERILL"
	8787. .9407
CDS	/codon_start=1
	/transl_table=1
	/product="nef protein"
	/protein_id="AAK08490.1"
	/db_xref="GI:12831143"
	/translation="WGGKWSKSSVIGPRAVERMRAEPADGVGAVSRLKHAIT SSNTANNAACAMLEADEEBEGVEPVYQVRLPMTYKKAADVLSHFLKEKGGLGLIH SORRODILDLMIYHOGYFPDMONITPGPVRYPLTFGMCYKLVAPVEDKVEANKGE NTSLSLHVSLSHGDDDPREYVLEWRPDSRLAFHNHVARLHREYFKNC"
	9076. .9709
LTR	/note="3' long terminal repeat"
	9529. .9626
repeat_region	

	polyA-signal	/note="R; 3' copy"
	misc_feature	9602..9607
		9710..11282
		/note="3' cellular genomic sequence"
	CDS	complement(12278..13138)
		/note="from pUC vector"
		/codon_start=1
		/transl_table=11
		/product="beta-lactamase"
		/protein_id="AAK08491.1"
		/db_xref="GI:12831144"
		/translation="MSIOHFRVALIPPPAFCLEVFAPHEPTLVVKDAEDOLGARVGY LELDNSGKILIESFRPEPERFMSTFVVLCGAALSRIDAQOEOIGRIIHSONDLVE YSPYTEKHLDGTGMIVRELCSNAITNMTNALITLITIGGRKELTAEIHMNDKVATLE DRMPELNELAIPINDERDPTMPVAMATLRKLIELLELTLSROOLDJENNEDKXAGVD LRSLAPGMFFIAODKSAGERSRCIIIALGPDGKPSRIVIYTTGSATDNDEHRQLINA EIGSLIKHW"
	misc_feature	13647..14824
		/note="5' cellular genomic sequence"
	BASE COUNT	4733 a 3022 c 3607 g 3462 t
	ORIGIN	
	Query Match	43.0%; Score 2676.2; DB 12; Length 14824;
	Best Local Similarity	99.9%; Pctd. No. 0;
	Matches 2678; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
Oy	1291 CCATGAGAGTGAAGAGAGAAGTAGTCACGACTTGCGAGATGGGGGTGANAATGGGGCACCA	1350
Db	6219 CAATGAGAGTGAAGAGAGAAGTAGTCACGACTTGCGAGATGGGGGTGANAATGGGGCACCA	6278
Oy	1331 TGCCTCTTGGCATATTGATGATCTGTGTAAGTCTTACAGAAAATTTGGGGTCAACGCTTATT	1410
Db	6279 TGCTCCTTGGCATATTGATGATCTGTGTAAGTCTTACAGAAAATTTGGGGTCAACGCTTATT	6338
Oy	1411 ATGGGGTACCCTGGTGGAGAGAACACACACCTTATTTTGGCATCAGATGCTTAAG	1470
Db	6339 ATGGGGTACCCTGGTGGAGAGAACACACACCTTATTTTGGCATCAGATGCTTAAG	6398
Oy	1471 CATATGATATACAGAGGTACATATATGTTTGGCCACACATCGCTGTACCCACACACCCA	1530
Db	6399 CATATGATATACAGAGGTACATATATGTTTGGCCACACATCGCTGTACCCACACACCCA	6458
Oy	1531 ACCCAACAAGATGTATTTGGTAAATGTGCAGAAAAATTTTAACTGTGAAAAATGACA	1590
Db	6459 ACCCAACAAGATGTATTTGGTAAATGTGCAGAAAAATTTTAACTGTGAAAAATGACA	6518
Oy	1591 TGTGTGAACAACATGCATGAGATATTAACAGTTTATGGGATCAAAGCCATGATGTG	1650
Db	6519 TGTGTGAACAACATGCATGAGATATTAACAGTTTATGGGATCAAAGCCATGATGTG	6578
Oy	1651 TAAATTAACCCCACCTCTGTGTGGTTTAAAGTGACAGTATTTGAAGTAAGTACTATA	1710
Db	6579 TAAATTAACCCCACCTCTGTGTGGTTTAAAGTGACAGTATTTGAAGTACTATA	6638
Oy	1711 CCAATAGTAGTAGGGGAGAAATGATTAATGAGAAAAGAGATATAAAAAATCGCTCTTCA	1770
Db	6639 CCAATAGTAGTAGGGGAGAAATGATTAATGAGAAAAGAGATATAAAAAATCGCTCTTCA	6698
Oy	1771 AATATCAGACAACATTAAGACATAAGGTGCAGAAAGAAATATGATCTTTTATATAACTTG	1830
Db	6699 AATATCAGACAACATTAAGACATAAGGTGCAGAAAGAAATATGATCTTTTATATAACTTG	6758
Oy	1831 AATATAGTACCAATGATATATACCGACTATAGGTGATATGTTGAACACCGTCACTCANTTA	1890
Db	6759 AATATAGTACCAATGATATATACCGACTATAGGTGATATGTTGAACACCGTCACTCANTTA	6818
Oy	1891 CACAGGCCCTGTCCAAAGGTATCTTTGAGCCAAATTCATATATGTGCCCCGGGTG	1950
Db	6819 CACAGGCCCTGTCCAAAGGTATCTTTGAGCCAAATTCATATATGTGCCCCGGGTG	6878
Oy	1951 GTTTTGGCATTTAATATGTAATAAAGAGCTTCAATGGAACAGGACCATGTACAAATG	2010
Db	6879 GTTTTGGCATTTAATATGTAATAAAGAGCTTCAATGGAACAGGACCATGTACAAATG	6938

QY	2011	TCAGCACAGTACATATGTCACATGGAAATCAGGCCAGTAGTATCAATCTCACTGCTGTTAA	2070
Db	6939	TCAGCACAGTACAAATGTACACATGGAAATCAGGCCAGTAGTATCAATCTCACTGCTGTTAA	6998
QY	2071	ATGGCAGCTGACAGAAAGACGTAGTAATATAGATCTGCCAATTTCCACAGACATGCTA	2130
Db	6999	ATGGCAGCTGACAGAAAGACGTAGTAATATAGATCTGCCAATTTCCACAGACATGCTA	7058
QY	2131	AAACCATATAGTACAGCTGAACACATCTGTAGAAATTTAATTTGTACAAGACCCAAACA	2190
Db	7059	AAACCATATAGTACAGCTGAACACATCTGTAGAAATTTAATTTGTACAAGACCCAAACA	7118
QY	2191	ATACAGAAAAAGATTCCTGATCCAGAGGGGACCAGAGAGCATTTGTTCATATAGGA	2250
Db	7119	ATACAGAAAAAGATTCCTGATCCAGAGGGGACCAGAGAGCATTTGTTCATATAGGA	7178
QY	2251	AAATGAGAAATATGAGACAGACACATTTTAACATTTAGTAGAGCAAAATGGAATGCCACTY	2310
Db	7179	AAATGAGAAATATGAGACAGACATTTTAACATTTAGTAGAGCAAAATGGAATGCCACTY	7238
QY	2311	TAAACCAATATAGCTAGCAAAATTTAGAGAAACAATTTGGAAATATATAAACATATCTTTA	2370
Db	7239	TAAACCAATATAGCTAGCAAAATTTAGAGAAACAATTTGGAAATATATAAACATATCTTTA	7298
QY	2371	AGCAATCCCTCAGAGAGGGGACCAGAAATTTGACGACAGCTTTTAATTTGTGGAGGGANAT	2430
Db	7299	AGCAATCCCTCAGAGAGGGGACCAGAAATTTGACGACAGCTTTTAATTTGTGGAGGGANAT	7358
QY	2431	TTTTCTACTGTATATCAACACACTGTTTAAATAGTACTTGCTTTAATATATGTACTTGGAGTA	2490
Db	7359	TTTTCTACTGTATATCAACACACTGTTTAAATAGTACTTGCTTTAATATGTACTTGGAGTA	7418
QY	2491	CTGAAGGGTCCAAATTAACACTGTAAGAGACTGTACACATACACTCCCATGACGAATTAAC	2550
Db	7419	CTGAAGGGTCCAAATTAACACTGTAAGAGACTGTACACATACACTCCCATGACGAATTAAC	7478
QY	2551	AATTTATTAACATGTGGCAGAGAGTAGGAAAAAGCAATGTATGCCCTCCATCAGTGGAC	2610
Db	7479	AATTTATTAACATGTGGCAGAGAGTAGGAAAAAGCAATGTATGCCCTCCATCAGTGGAC	7538
QY	2611	AAATTAGAGTTTCATCAATATATTTACTGGGGCTGTATTAACAAGATGCTGTAAATACA	2670
Db	7539	AAATTAGAGTTTCATCAATATATTTACTGGGGCTGTATTAACAAGATGCTGTAAATACA	7598
QY	2671	ACAATGGGTCGAGATCTTCAGACCTCGAGAGAGCGATATAGAGGACAAATTTGAGAAAGTG	2730
Db	7599	ACAATGGGTCGAGATCTTCAGACCTCGAGAGAGCGATATAGAGGACAAATTTGAGAAAGTG	7658
QY	2731	AATTTATTAATATTAAGTAGTAAAAAATTTGAACCTTAGAGTAGACCCACCAAGGCAA	2790
Db	7659	AATTTATTAATATTAAGTAGTAAAAAATTTGAACCTTAGAGTAGACCCACCAAGGCAA	7718
QY	2791	AGAGAAAGTGGTGGCAGAGAGAAAAAAGAGAGTGGGAAATAGAGACTTTGTTCCTGGGT	2850
Db	7719	AGAGAAAGTGGTGGCAGAGAGAAAAAAGAGAGTGGGAAATAGAGACTTTGTTCCTGGGT	7778
QY	2851	TCTTGGAGACAGCAGAAACACTATAGGCTGCACGCTCAATGACAGCTGACGGTAGAGGCA	2910
Db	7779	TCTTGGAGACAGCAGAAACACTATAGGCTGCACGCTCAATGACAGCTGACGGTAGAGGCA	7838
QY	2911	GACATTTATTTGCTGTATATATAGTGCAGCAGCAAGAACATTTTGCTGAGGGCTATTGAAGCGC	2970
Db	7839	GACATTTATTTGCTGTATATATAGTGCAGCAGCAAGAACATTTTGCTGAGGGCTATTGAAGCGC	7898
QY	2971	AACAGCATCTGTTGCAACTCAGAGCTGGGCAATCAACACACTCCAGGCAAGAAATCTCGG	3030
Db	7899	AACAGCATCTGTTGCAACTCAGAGCTGGGCAATCAACACACTCCAGGCAAGAAATCTCGG	7958
QY	3031	CTGTGGAAAGATACCTAAAGATCAACAGCTCTGGGGATTTTGGGGTTGCTCTGTGAAAC	3090
Db	7959	CTGTGGAAAGATACCTAAAGATCAACAGCTCTGGGGATTTTGGGGTTGCTCTGTGAAAC	8018

QY 3091 TCATTTGCACCACTGCTGCTTGGAAATGCTAGTGGAGTAAATATCTCTGGAACAGA 3150
 |||||||
 Db 8019 TCATTTGCACCACTGCTGCTTGGAAATGCTAGTGGAGTAAATATCTCTGGAACAGA 8078
 |||||||
 QY 3151 TTTGGAAATACATGACCTGATGAGTGGACAGAGAAATTAACAATTAACAACAGCTTAA 3210
 |||||||
 Db 8079 TTTGGAAATACATGACCTGATGAGTGGACAGAGAAATTAACAATTAACAACAGCTTAA 8138
 |||||||
 QY 3211 TACACTGCTTAATTAAGAAATCGCAAAACCGAACAAAGAAATGAATGAATATATGG 3270
 |||||||
 Db 8139 TACACTGCTTAATTAAGAAATCGCAAAACCGAACAAAGAAATGAATGAATATATGG 8198
 |||||||
 QY 3271 AATTAGATAAATGGGCAAGTTTGTGAATGGTTTAAACATAACAATTCGCTGTGTATA 3330
 |||||||
 Db 8199 AATTAGATAAATGGGCAAGTTTGTGAATGGTTTAAACATAACAATTCGCTGTGTATA 8258
 |||||||
 QY 3331 TAAATTTATCATATATGATAGTAGAGGCTTGTAGTGTAAAGATTTTTCGTGTAC 3390
 |||||||
 Db 8259 TAAATTTATCATATATGATAGTAGAGGCTTGTAGTGTAAAGATTTTTCGTGTAC 8318
 |||||||
 QY 3391 TTTCTATAGTGAATAGATTTAGGCGAGGATATTCACCTATTCCTTTCAGACCCACCTCC 3450
 |||||||
 Db 8319 TTTCTATAGTGAATAGATTTAGGCGAGGATATTCACCTATTCCTTTCAGACCCACCTCC 8378
 |||||||
 QY 3451 CAATCCGAGGGGAGCCGACAGGCCCGAAGAAATAGAAAGAAAGCTGGAGAGAGACA 3510
 |||||||
 Db 8379 CAATCCGAGGGGAGCCGACAGGCCCGAAGAAATAGAAAGAAAGCTGGAGAGAGACA 8438
 |||||||
 QY 3511 GAGACAGATCCATTCGATTAAGTGAACGATTCCTTAGCCTTATCTGGGACGATCTGCCGA 3570
 |||||||
 Db 8439 GAGACAGATCCATTCGATTAAGTGAACGATTCCTTAGCCTTATCTGGGACGATCTGCCGA 8498
 |||||||
 QY 3571 GCTGTGCTCTTCACGATACACCGCTTGAGAGACTTACTCTTGTATTAACGAGAGATTG 3630
 |||||||
 Db 8499 GCTGTGCTCTTCACGATACACCGCTTGAGAGACTTACTCTTGTATTAACGAGAGATTG 8558
 |||||||
 QY 3631 TGGAACTTCGAGGACGAGGGGGTGGGAAGCCCTCAATATTGGTGAATCTCTACAGT 3690
 |||||||
 Db 8559 TGGAACTTCGAGGACGAGGGGGTGGGAAGCCCTCAATATTGGTGAATCTCTACAGT 8618
 |||||||
 QY 3691 AATTGAGTCAGAGAACTAAGAAATAGTGTCTTAATCTCTCAATGCCACAGCAATAGAG 3750
 |||||||
 Db 8619 AATTGAGTCAGAGAACTAAGAAATAGTGTCTTAATCTCTCAATGCCACAGCAATAGAG 8678
 |||||||
 QY 3751 TAGCTGAGGGACAGATAGGGTTATAGAAATTAACAAGACGTTATAGACTATTCGCC 3810
 |||||||
 Db 8679 TAGCTGAGGGACAGATAGGGTTATAGAAATTAACAAGACGTTATAGACTATTCGCC 8738
 |||||||
 QY 3811 ACATACCTAGAAAGATATAGACAGGCTTGAAGGATTTTCTATTAAGATGGGTGGCAAG 3870
 |||||||
 Db 8739 ACATACCTAGAAAGATATAGACAGGCTTGAAGGATTTTCTATTAAGATGGGTGGCAAG 8798
 |||||||
 QY 3871 TGTGTCAAAAAGTATGATTTGATGATGCGCTCTGTAAAGGAAAGATAGAGAGAGCTGAG 3930
 |||||||
 Db 8799 TGTGTCAAAAAGTATGATTTGATGATGCGCTCTGTAAAGGAAAGATAGAGAGAGCTGAG 8858
 |||||||
 QY 3931 CCAGCAGCAGATGGGGTGGAGACGATATCTCGAGATCTTAGA 3971
 |||||||
 Db 8859 CCAGCAGCAGATGGGGTGGAGACGATATCTCGAGACCTAGA 8899
 |||||||

RESULT 7
 HIV1U26942
 LOCUS HIV1U26942 9000 bp DNA linear VRL 15-NOV-1995
 DEFINITION Human immunodeficiency virus clone pNL4-3, subclone 4.20.
 ACCESSION U26942
 VERSION U26942.1 GI:902798
 KEYWORDS
 SOURCE Human immunodeficiency virus 1 (HIV-1)
 ORGANISM Human immunodeficiency virus 1
 Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
 lentivirus group.
 REFERENCE 1 (bases 1 to 9000)

AUTHORS Salminen,M.O., Koch,C., Sanders-Buell,E., Ehrenberg,P.K.,
 Michael,N.L., Carr,J.K., Burke,D.S. and McCutchan,F.E.
 TITLE Recovery of virtually full-length HIV-1 provirus of diverse
 subtypes from primary virus cultures using the polymerase chain
 reaction
 JOURNAL Virology 213 (1), 80-86 (1995)
 MEDLINE 96036482
 PUBMED 7483282
 REFERENCE 2 (bases 1 to 9000)
 AUTHORS Adachi,A., Gendelman,H.E., Koenig,S., Folks,T., Willey,R.,
 Rabson,A., and Martin,M.A.
 TITLE Production of acquired immunodeficiency syndrome-associated
 retrovirus in human and nonhuman cells transfected with an
 infectious molecular clone
 JOURNAL J. Virol. 59 (2), 284-291 (1986)
 MEDLINE 86281827
 PUBMED 3016298
 REFERENCE 3 (bases 1 to 9000)
 AUTHORS Salminen,M.S.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAY-1995) Mika S. Salminen, Program area 1, Henry M.
 Jackson Foundation, 1600 East Gude Drive, Rockville, MD 20850, USA
 FEATURES
 source
 1..9000
 /organism="Human immunodeficiency virus 1"
 /mol_type="genomic DNA"
 /strain="NL4-3"
 /db_xref="taxon:11676"
 /note="Long PCR amplified, re-cloned 9kb fragment from
 infectious HIV-1 clone pNL4-3, subclone 4.20. Insert
 sequence only, vector is pMOT. Cloned into invitrogen
 vector pMOTa from extended PCR amplification product using
 original pNL4-3 molecular clone as template. Only 17
 mutations compared to original published sequence (GenBank
 Accession Number M19921) were observed."
 <1..13
 /note="5' LTR"
 168..1670
 /gene="gag"
 168..1670
 /gene="gag"
 /note="matrix (p17), capsid (p24) and core (p7, p6)
 proteins"
 /codon_start=1
 /product="Gag polyprotein precursor"
 /protein_id="AAB60571.1"
 /db_xref="GI:902798"
 /translation="MGARASVLSGCELDKWEKRLPRGCKKQYKRIHWASPELEP
 AVNPGLETSBGCNIIIGOLPISLOTGSEELRSILNTIAYLYCHQIDVXDKREALD
 KIEEONKSKRKRAQOAAADTGNNSQVSONPIYVNLQGOAHQISPTLNAVKKVE
 EKAPSEVIMPEFSALSEGATPODNLQMLNVGGQAMOMLKEITINEEAEMWLRHFV
 HAGPIAGOMKEPBGSDIAGTSTLQELQMLNVTGQAEVKNMTEFLIVONANPCKTIL
 YSPSIDIRQGPKEPPRDVDFRYKTLRLQEAQOEKNNMTEFLIVONANPCKTIL
 KALGPAGTLEEMTAQCOGVGPGHKARVLAEMASQVNPATIMIOKGNPNOKRTVC
 PNCKEESHIAKNCAAPRRKKGCMKCGKGRHOKDKTEQANFLKTIKPSHKGRENFLQ
 SRPEPTAPPESEFRGEETTPSORPEIDKELYPIASLSLSLSDPSQ"
 1176
 /gene="gag"
 /citation=[2]
 /replace="t"
 1463..4474
 /gene="pol"
 <1463..4474
 /gene="pol"
 /note="protease, RT, RNase H, Integrase; Translated by -1
 frameshift at TT codon at end of gag p7"
 /codon_start=1
 /product="POL polyprotein precursor"
 /protein_id="AAB60572.1"
 /db_xref="GI:902800"
 /translation="EPRERDLAPQGRKARESSRQTRANSPTRRELQVWGRDNNSLSFA
 GADRQGVSPFSFQITILMORPLVITIKGGQLEKALDITGADDTVLEEMNIPGRKKPM
 IGGIGFIRKQIDQILIEICGHRKAITGVGPTPVNITGRNLTITIGCTLNPPIPSPI

OY	1651	TAATAATTACCCCACTGCTGTGTAGTTTAAAGTCAGTACTGATTTGAGATGATATCTATA	1710
Db	5957	TAAATTTAACCCCACTGCTGTGTAGTTTAAAGTCAGTACTGATTTGAGATGATATCTATA	6016
OY	1711	CCAATAGTAGAGGAGGAGAAATGAAATGAGAAAGGAGATATAAAAACATGCTCTTCA	1770
Db	6017	CCAATAGTAGAGGAGGAGAAATGAAATGAGAAAGGAGATATAAAAACATGCTCTTCA	6076
OY	1771	ATATGACACAGCATTAAGATTAAGGTGAGAAAAGAAATATGATCTTTTATTAACCTG	1830
Db	6077	ATATGACACAGCATTAAGATTAAGGTGAGAAAAGAAATATGATCTTTTATTAACCTG	6136
OY	1831	ATATAGTAGCAATGATTAATACCGATATAGCTTGATTAAGCTTGTAACCTCATCTTA	1890
Db	6137	ATATAGTAGCAATGATTAATACCGATATAGCTTGATTAAGCTTGTAACCTCATCTTA	6196
OY	1891	CACAGGCCCTGCCAAAGGATATCTTTGAGCCAAATTCATTCATTAATGTGCCCCGGCTG	1950
Db	6197	CACAGGCCCTGCCAAAGGATATCTTTGAGCCAAATTCATTCATTAATGTGCCCCGGCTG	6256
OY	1951	GTTTTGGATTTCTAAATGTAAATAAAGAGTTCATGGAACAGACCATGTACAATG	2010
Db	6257	GTTTTGGATTTCTAAATGTAAATAAAGAGTTCATGGAACAGACCATGTACAATG	6316
OY	2011	TCACACACGTATCATGTAACATGGAATCAGGCCAGTAGTAATCACTCACTGCTGTAA	2070
Db	6317	TCACACACGTATCATGTAACATGGAATCAGGCCAGTAGTAATCACTCACTGCTGTAA	6376
OY	2071	ATGGCAGTCTGCAGAAAGAAATGTAGTAATTAATCATCTGCCAATTCACAGACAACTGA	2130
Db	6377	ATGGCAGTCTGCAGAAAGAAATGTAGTAATTAATCATCTGCCAATTCACAGACAACTGA	6436
OY	2131	AAACCATTAATGTAACAGCTGAACACATCTGTAGAAATTAATTTAGCAAGACCACAACA	2190
Db	6437	AAACCATTAATGTAACAGCTGAACACATCTGTAGAAATTAATTTAGCAAGACCACAACA	6496
OY	2191	ATTACAGAAAAGATTCGGTATCCAGAGGGGACACAGGAGAGCAATTTGTTACATTAAGAA	2250
Db	6497	ATTACAGAAAAGATTCGGTATCCAGAGGGGACACAGGAGAGCAATTTGTTACATTAAGAA	6556
OY	2251	AAATAGAAATATGAGACACACATTTGTACATTTAGTAGACAAATGCAATGCCACTT	2310
Db	6557	AAATAGAAATATGAGACACACATTTGTACATTTAGTAGAGCAAAATGGAATGCCACTT	6616
OY	2311	TAAACAGATAGCTAGCAAAATTTAAGAGAACATTTGGAATTAATTAACATAATCTTTA	2370
Db	6617	TAAACAGATAGCTAGCAAAATTTAAGAGAACATTTGGAATTAATTAACATAATCTTTA	6676
OY	2371	AGCAATCCTCAGGAGGGGACCCCAAAATTTGTAACGCACAGTTTAATTTGAGGGGAAT	2430
Db	6677	AGCAATCCTCAGGAGGGGACCCCAAAATTTGTAACGCACAGTTTAATTTGAGGGGAAT	6736
OY	2431	TTTTCTACTGTAATTCACACACTGTTTAAATATCTACTGGTTTAATAGTACTGGAATA	2490
Db	6737	TTTTCTACTGTAATTCACACACTGTTTAAATATCTACTGGTTTAATAGTACTGGAATA	6796
OY	2491	CTGAAGGCTCAAAATACACTGGAAGAGTGACACAATCACACTCCCATGCAAGATAAATAC	2550
Db	6797	CTGAAGGCTCAAAATACACTGGAAGAGTGACACAATCACACTCCCATGCAAGATAAATAC	6856
OY	2551	AATTTTATTAACATGTGCGAGGAAGTAGAAAAAGCAATGATATGCCCTCCCATCAGTGAC	2610
Db	6857	AATTTTATTAACATGTGCGAGGAAGTAGAAAAAGCAATGATATGCCCTCCCATCAGTGAC	6916
OY	2611	AAATTAAGTTCATCAAAATTAATTAATCTGGGCTGCTATTAACAAGATGGGTGTAATACA	2670
Db	6917	AAATTAAGTTCATCAAAATTAATTAATCTGGGCTGCTATTAACAAGATGGGTGTAATACA	6976
OY	2671	ACAATGGGCTCGAATCTTCAGACCTGAGAGGCGAATGAGAGCAATTTGGAGAGTG	2730
Db	6977	ACAATGGGCTCGAATCTTCAGACCTGAGAGGCGAATGAGAGCAATTTGGAGAGTG	7036
OY	2731	AATTAATATAATATAAGTAGATAAATAATTTGACCATTAAGAGTAGCAACCCACAGGACAA	2790

D	b		7037	AATPATATAATAAAGTAGTAAAAATTGAACCTTAGAGAGCACCACCAAGCAA	7096
O	y		2791	AGAGAAGAGTGTCGACAGAGAAAAAAGACACTGGGAATAGAGCTTTGTCTTGGGT	2850
D	b		7097	AGAGAAAGTGTCGACAGAGAAAAAAGACACTGGGAATAGAGCTTTGTCTTGGGT	7156
O	y		2851	TCTTGGAGCAGCAGAGAACCTATGCGGTGCACGTCAATGACGCTGACGGTACAGCCA	2910
D	b		7157	TCTTGGAGCAGCAGAGAACCTATGCGGTGCACGTCAATGACGCTGACGGTACAGCCA	7216
O	y		2911	GACAATATTTGCTGATATATGTCGACAGCAACAATTTGCTGAGGGCTATTGAGGGCC	2970
D	b		7217	GACAATATTTGCTGATATATGTCGACAGCAACAATTTGCTGAGGGCTATTGAGGGCC	7276
O	y		2971	AACAGCATCTGTGGCACTCACAGTCTGGGGCATCAAAAGCTCCAGCAGCAAACTCTGG	3030
D	b		7277	AACAGCATCTGTGGCACTCACAGTCTGGGGCATCAAAAGCTCCAGCAGCAAACTCTGG	7336
O	y		3031	CTGTGGAAGATACCTTAAGAGATCAACAGCTCTCTGGGGATTGGGGTTCTCTGGAANA	3090
D	b		7337	CTGTGGAAGATACCTTAAGAGATCAACAGCTCTCTGGGGATTGGGGTTCTCTGGAANA	7396
O	y		3091	TCAATTTCCACCACTGCTGCTGCTTGGAAATGCTATGTTGAGATTAATCTCTGGAACGA	3150
D	b		7397	TCAATTTCCACCACTGCTGCTGCTTGGAAATGCTATGTTGAGATTAATCTCTGGAACGA	7456
O	y		3151	TTTGAATTAACATGACCTGATGAGAGTGGAGCAGAGAAATTAACAATTAACACACTTA	3210
D	b		7457	TTTGAATTAACATGACCTGATGAGAGTGGAGCAGAGAAATTAACAATTAACACACTTA	7516
O	y		3211	TACACTCCTTAATTAATGAAGAAATGCCAAAACAGCAGAGAAAGATGAACAGAAATTAATGG	3270
D	b		7517	TACACTCCTTAATTAATGAAGAAATGCCAAAACAGCAGAGAAAGATGAACAGAAATTAATGG	7576
O	y		3271	AATTAAGTAATATGGGGCAAGTTTGGAAATGTTTTAACTPAACAAATTTGGCTGGCTAPA	3330
D	b		7577	AATTAAGTAATATGGGGCAAGTTTGGAAATGTTTTAACTPAACAAATTTGGCTGGCTAPA	7636
O	y		3331	TAAATATTTTCATPATATGATAGTAGAGAGCTTGGTAGTTTAAGAAATAGTTTTGCTGTAC	3390
D	b		7637	TAAATATTTTCATPATATGATAGTAGAGAGCTTGGTAGTTTAAGAAATAGTTTTGCTGTAC	7696
O	y		3391	TTTCTATATGATATAGAGTTAGGCGAGGATATTCAACATTATTCGTTTCAGACCCACTCC	3450
D	b		7697	TTTCTATATGATATAGAGTTAGGCGAGGATATTCAACATTATTCGTTTCAGACCCACTCC	7756
O	y		3451	CAATCCGAGAGGGGACCCGACAGGCCCGGAAGGATTAAGAAAGGTTGGAGAGAGACA	3510
D	b		7757	CAATCCGAGAGGGGACCCGACAGGCCCGGAAGGATTAAGAAAGGTTGGAGAGAGACA	7816
O	y		3511	GAGACATATCATTTTCATTTAGTGAACGATCTTTACACTTTATCTTGGAGCATCTGGCGA	3570
D	b		7817	GAGACATATCATTTTCATTTAGTGAACGATCTTTACACTTTATCTTGGAGCATCTGGCGA	7876
O	y		3571	GCCGTGTCCTTTCAGCTACACCGCTTGAGAGACTTTACTCTTGATTTGAACGAGATTTG	3630
D	b		7877	GCCGTGTCCTTTCAGCTACACCGCTTGAGAGACTTTACTCTTGATTTGAACGAGATTTG	7936
O	y		3631	TGGAACCTTCTGGAGCAGGGGTTGGGAAGCCCTCAATATTGTTGGTGAATCTCTACAGT	3690
D	b		7937	TGGAACCTTCTGGAGCAGGGGTTGGGAAGCCCTCAATATTGTTGGTGAATCTCTACAGT	7996
O	y		3691	ATTGGAATTCAGGAACTTAAGAAATATGATGCTTTAACTTGTCAATGCGCAGCATAAGAG	3750
D	b		7997	ATTGGAATTCAGGAACTTAAGAAATATGATGCTTTAACTTGTCAATGCGCAGCATAAGAG	8056
O	y		3751	TAGCTGAGGGGACAGATAGGGTTATAGAGATTTATCAACACTTATAGAGCTATTTCGCC	3810
D	b		8057	TAGCTGAGGGGACAGATAGGGTTATAGAGATTTATCAACACTTATAGAGCTATTTCGCC	8116
O	y		3811	ACATACCTTAGAAGATTAAGACAGGGCTTGGAAAGATTTTGCCTATTAAGATGGGTGGCA	3870

Db 8117 ACATACCTAGAAAGATAAGACAGGCGCTTGAAAGGATTTCCTATATAGATGGTGGCAAG 8176
Oy 3871 TGTCAAAAAGTAGTGTATGGATGGCTGCTGTAAGGAGAAAGATGACGAGCTGAG 3930
Db 8177 TGGTCAAAAAGTAGTGTATGGATGGCTGCTGTAAGGAGAAAGATGACGAGCTGAG 8236
Oy 3931 CCACGACGAGATGGGTGGAGCAGTATCTCGAGATCTAGA 3971
Db 8237 CCACGACGAGATGGGTGGAGCAGTATCTCGAGACCTAGA 8277

RESULT 8
HIV2132 9754 bp DNA linear VRL 25-DEC-2002
LOCUS Human immunodeficiency virus 1 DNA for Gag, Pol, Vif, Vpr, Rev,
DEFINITION Tat, Env, Vpu.
ACCESSION D86069.1 GI:1398964
VERSION D86069.1 GI:1398964
KEYWORDS Vpu; Env; Tat; Rev; Vpr; Vif; Pol; Gag.
SOURCE Human immunodeficiency virus 1 (HIV-1)
ORGANISM Human immunodeficiency virus 1
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
Lentivirus group.
REFERENCE 1
AUTHORS Cloyd,M.W. and Moore,B.E.
TITLE Spectrum of biological properties of human immunodeficiency virus
(HIV-1) isolates
JOURNAL Virology 174 (1), 103-116 (1990)
MEDLINE 90101366
PUBMED 1688473
REFERENCE 2 (bases 1 to 9754)
AUTHORS Iwatani,Y.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 9754)
AUTHORS Iwatani,Y.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-1996) Yasumasa Iwatani, Institute for Virus
Research, Kyoto University, Laboratory of Gene Analysis; Sakyo-ku,
Shogoin-Kawahara-machi, Kyoto, Kyoto 606-01, Japan
(E-mail:y.iwatani@virus.kyoto-u.ac.jp, Tel:075-751-3996,
Fax:075-751-3995)

FEATURES
source
Location/Qualifiers
1..9754
/organism="Human immunodeficiency virus 1"
/mol_type="genomic DNA"
/strain="pm213"
/isolate="213"
/db_xref="taxon:11676"
1..636
790..2328
/gene="gag"
790..2328
/gene="gag"
/codon_start=1
/product="Gag"
/protein_id="BA12996.1"
/db_xref="GI:1398965"
/translation="MGARASVLSGGKLDNRKIKLRPGCKKKYKLIKHIWASRLER
AVRGLTETSEGCQIILGLOPISLQTSSEERSLYNTVATLYCHOKRIETKDTLEAD
KIEEONKSKKKAQAAANTGHSSOVSQNPPIYONIGOMVHOAISPTLNAWVVE
EKASPEVIMFSAISGATPODINTMINTGGHOAMOMLKETINBEAAEMDVHVV
HAGPIAGOMREPGSDIAGTSTLOEIGMTNNPPIPVGEIKRWIILGLNKIVRA
YSPSTIDIRGPREPRDYDRYKTLRAOASOEYKMMTEFTLLYONANPDCKTIL
KALGPAATLEEMNACOGVSGPGKARVLAEMQVNTATIMKRGFRORAKMYC
PNCKEHTANCRAPRKKGWCCKGSGHOKDCTERQANFLGKIMPSYKGRPNFLQ
SRPEPTAPPLQSRPEPTAPPEESFSGVEATTPPOKPEIDKELVPLTSLRSLFGND
PS50"

gene 2085..5132
/gene="pol"
<2085..5132
/gene="pol"
/codon_start=1
/product="Pol"

CDS

/protein_id="BA12997.1"
/db_xref="GI:1398966"
/translation="EPREDLAFLOGKAREEESBOTRANSPITISEOTRANSPFREIO
VMGKNNSPSAGADROGTSFENRPOITIMORPLATYTIKIGOLKEALLDGTADTYLE
EWSLPGRRKPRMIGIGGFIKRVOTDILIEICGKAIGIVLVPTPVNIIIRNLTLQ
IGCTLNFPISPIETVPYKLGKMDGPKVOMPLIEEKIKALVEICTEMEKKGISKJG
PENNYNTPVEFAIKKSDSTKMRKLVDPFMRITODFNEVOLGIEPHGAKKKSVTVL
DVGAYTSVPLDEDFRKYTATTISINNEPFGIYOVNVLDGKSGAPAIQSSMTLI
LEPRKONPDIVYIOYMDLYVGSDELIGORHKEIEELROHLIRMGITTPDKRKKRP
PFLMGYLEHPDKNTVOPVYLPKDSMTVNDIQVLGKRLMASOITPRTKROJCKLL
RGTALETEVPLTEAELELAENRELIKEPHGVYDPSKDLAEIQKGGQMTYOL
YOEPPKMLKGTGVARTRGANTNDVKOLEAVOKITTESIYVMGTPFKPLIEQKTE
TMWTMEYQATWIPMEPEVNPPLVKLWYOLEKEPIVGAEEFYVDGASRETKKAGY
VTNRGROKVATLTDTTNOKTELOALIVALODSGEIVNVLDGKSGAPAIQSSMTLS
ELVNOIIEQLKREKVYLAWPRAHKGAGNEQVYKIVSAGIRKILFTIDGKADDEHE
KYSNMRAASDFNLPVVAKEIVASCDKCOLKEAMHGOVDCSPGIMQDCTHLEK
VILVAHVASGYLEAEVYIPAEITQGETAYFLKTLKLRMPVKTIHDNSNFTSTVKA
CWMAGIKOEFGIIPYNPQSGVYESMNEELKKIIQGVDOAEHLTKDAVOMAVFIHNF
KGGIGYSAGEIYDIIATDIOTELQKOLTKIONFPRYVHDSRDLPMKGPALKMKG
EGAVVIDNSDIKVPYPRKAKIIRDYKQAMGDCVASRODED"
5077..5655
/gene="vif"
5077..5655
/gene="vif"
/codon_start=1
/product="Vif"
/protein_id="BA12998.1"
/db_xref="GI:1398967"
/translation="MENRMQVMIYOYDRMRIRTWKSILKHHMYVSGKRGWEYRHHY
ESPHRISSEVHIPLGDARLVIYTWGLHGERMHGQVSIEMRRKRRSTOYDPEL
ADOLILHYEPDCFSDSAIRKALLGHIVSPREYQAGHNKVGSLQYLAALITPKKYK
PLPISVTKLTEDRNNKPKQTKGHRGSHTMNGH"
5595..5708
/gene="vpr"
5595..5708
/gene="vpr"
/codon_start=1
/product="Vpr"
/protein_id="BA12999.1"
/db_xref="GI:1398968"
/translation="MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHPPRI"
/join(5866..6080,8414..8459)
/gene="tat"
/join(5866..6080,8414..8459)
/gene="tat"
/codon_start=1
/product="Tat"
/protein_id="BA13000.1"
/db_xref="GI:1398969"
/translation="MEPVDPRLPEMKHGSQPTACTNCYCKKCFHQCVCFTIKALG
ISYGRKKRQRRRAHONSQTHQVLSKOPTSQPRGDPFGPRE"
/join(6005..6080,8414..8688)
/gene="rev"
/join(6005..6080,8414..8688)
/gene="rev"
/codon_start=1
/product="Rev"
/protein_id="BA13001.1"
/db_xref="GI:1398970"
/translation="MAGRSQGSDELIRTVRLIKELYQSNNPPNPEGTROARRRRR
WRERORQIHSISERILSTYLGRSADAVPLQPLPERLLTDCNEDCGTSGDGVGSP
LVESPTVLESTKE"
6097..6345
/gene="vpu"
6097..6345
/gene="vpu"
/codon_start=1
/product="Vpu"
/protein_id="BA13002.1"
/db_xref="GI:1398971"
/translation="TQPIPIVTIVALVAIIITAIIVMSIVIIIEYRKILRORRIDRLID
RLIERADSGNESEGEISALVEMGVENHAPMDVDDL"
6260..8830
gene

Query Match	Best Local Similarity	41.9%	Score 2610.6;	DB 14;	Length 9754;			
Matches 26477;	Conservative	0;	Mismatches	34;	Indels	6;	Gaps	1;
1291	CCATGAGAGTGAAGAGAGAGTATCAGCAGCTGTGGAGATGGGGGTGGAATGGGGACCA	1350						
6258	CAATGAGAGTGAAGAGAGAGATATCAGCAGCTGTGGAGATGGGGGTGGAAGATGGGGACCA	6317						
1331	TGCTCCTTGGGATTTGATGATCTGTAGTGTCTACAGAAAAATTTGGGTACCGCTTAAT	1410						
6318	TGCTCCTTGGGATTTGATGATCTGTAGTGTCTACAGAAAAATTTGGGTACAGAGCTTAAT	6377						
1411	ATGGGGTACCGTGTGGGAAGAACCAACCACTAATTTTGGCATTCAGATGCTTAAG	1470						
6378	ATGGGGTACCGTGTGGGAAGAACCAACCACTAATTTTGGCATTCAGATGCTTAAG	6437						
1471	CATATGATACAGAGTACATTAATTTTGGGACACATGCTGTGTACCCACAGACCCCA	1530						
6438	CATATGATACAGAGTACATTAATTTTGGGACACACATGCTGTGTACCCACAGACCCCA	6497						
1531	ACCACAGAGAGTATGGTAAATGTGACAGAAAAATTTTACATGTGAAAAATGACA	1590						
6498	ACCACAGAGAGTATGGTAAATGTGACAGAAAAATTTTACATGTGAAAAATGACA	6557						
1591	TGATGAGAACAGATGCATGAGATATTAATCACTTATGGGATCAAGCCCTAAGCCATGTG	1650						
6558	TGATGAGAACAGATGCATGAGATATTAATCACTTATGGGATCAAGCCCTAAGCCATGTG	6617						
1651	TAAATTAACCCCACTGTGTGTTTAAAGTGCACGTGATTTGAAGAAGATCTATA	1710						
6618	TAAATTAACCCCACTGTGTGTTTAAAGTGCACGTGATTTGAAGAAGATCTATA	6677						
1711	CCAAATAGTAGGAGGAGATGATTAATGGAGAAAGAGAGATTAAGAACTGCTCTTCA	1770						
6678	CCAAATAGTAGGAGGAGATGATTAATGGAGAAAGAGAGATTAAGAACTGCTCTTCA	6737						
1771	ATATCAGCACAGCATTAAGAGATTAAGTGCAGAAAGATATGATCTCTTTTATTAACCTG	1830						
6738	ATATCAGCACAGCATTAAGAGATTAAGTGCAGAAAGATATGATCTCTTTTATTAACCTG	6797						
1831	ATATAGTACCAATAGATTA-----TACAGGCTATAGGTGATAGTGTGAACCCCTAG	1884						
6798	ATATAGTACCAATAGATAGATATCTACACAGCTATTAAGTGTGAACAGATTGTACACCTAG	6857						
1885	TCATTTACAGAGCCCTGTCCAAAGTATCTTTGAGCAATTTCCATTCATTAATGTGCC	1944						

Db	6858	TCATTACACAGCGCTGTCCAAAGGTATTCCTTGGACCAATTCCTCATCATTTATTTGTGCC	6917
QY	1945	CGCGTGTGTTTTGCCATTTCTTAATAATGTAAATTAAGACGTTTCATATGAACAGAGCATGTGA	2004
Db	6918	CGCGTGTGTTTTGCCATTTCTTAATAATGTAAATTAAGACGTTTCATATGAACAGAGCATGTGA	6977
QY	2005	CAAAATGTCAGCAGCATATGTACACATGGAATCAGGCCAGTAGATATCAATCTCACTGC	2064
Db	6978	CAAAATGTCAGCAGCATATGTACACATGGAATGGAATGAGCCAGTAGATATCAATCTCACTGC	7037
QY	2065	TGTTAAATGAGAGCTAGCAGAAAGATGTAGTAATTAGTCTGCCAATTTCCACAGACA	2124
Db	7038	TGTTAAATGAGAGCTAGCAGAAAGATGTAGTAATTAGTCTGCCAATTTCCACAGACA	7097
QY	2125	ATGCTAAATACCATATAGTAGACGTGAACATCTGTGAAATTAATTTGTACAGACCCA	2184
Db	7098	ATGCTAAATACCATATAGTAGACGTGAACATCTGTGAAATTAATTTGTACAGACCCA	7157
QY	2185	ACAACAAATACAGAAAAAGTATCCGTAATCCAGAGGGGACAGGAGAGCATTTTGTTCAAA	2244
Db	7158	ACAACAAATACAGAAAAAGTATCCGTAATCCAGAGGGGACAGGAGAGCATTTTGTTCAAA	7217
QY	2245	TAGGAAAAATATGGAATATGAGACACACATTTGAACATTTAGTAGACCAAAATGCAATG	2304
Db	7218	TAGGAAAAATATGGAATATGAGACACACATTTGAACATTTAGTAGACCAAAATGCAATG	7277
QY	2305	CCACTTTAAACAGATAGCTAGCAAAATTAAGAAACATTTGGAAATTAATTAACATTA	2364
Db	7278	CCACTTTAAACAGATAGCTAGCAAAATTAAGAAACATTTGGAAATTAATTAACATTA	7337
QY	2365	TCTTTAAACAATCCTCAGAGAGGGGACCCGAAATTTGTAACGACACAGTTTTAATTTGGAG	2424
Db	7338	TCTTTAAACAATCCTCAGAGAGGGGACCCGAAATTTGTAACGACACAGTTTTAATTTGGAG	7397
QY	2425	GCGAATTTTCTACTGTAATTCACACACACTGTTTAATAGTACTTGGTTTAATATGTACTT	2484
Db	7398	GCGAATTTTCTACTGTAATTCACACACACTGTTTAATAGTACTTGGTTTAATATGTACTT	7457
QY	2485	GGAGACTGAAGGGGCAATTAACACTGAAGAGACTGACACAAATCACCTCCCATGACGAA	2544
Db	7458	GGAGACTGAAGGGGCAATTAACACTGAAGAGACTGACACAAATCACCTCCCATGACGAA	7517
QY	2545	TAAACAATTTATTAACATGTGGCAGAGAAATGAGAAAAAGCAATGTATGCCCTCCCATCA	2604
Db	7518	TAAACAATTTATTAACATGTGGCAGAGAAATGAGAAAAAGCAATGTATGCCCTCCCATCA	7577
QY	2605	GTGGCAAAATTAGATGTTTCATCAAAATATTACTGSGGCTGATTTAACAAGAGATGGTGA	2664
Db	7578	GTGGCAAAATTAGATGTTTCATCAAAATATTACTGSGGCTGATTTAACAAGAGATGGTGA	7637
QY	2665	ATAACAAATATGGGTCGAGATCTTTCAGACTGGAGGAGCGGATATGAGGACAAATGGGA	2724
Db	7638	ATAACAAATATGGGTCGAGATCTTTCAGACTGGAGGAGCGGATATGAGGACAAATGGGA	7697
QY	2725	GAAATGAAATTTATTAATATTAATTAAGTAGTAAATAATTGAACCATTTGAGATAGCACCCACA	2784
Db	7698	GAAATGAAATTTATTAATATTAATTAAGTAGTAAATAATTGAACCATTTGAGATAGCACCCACA	7757
QY	2785	AGGCAAAAGAGAAGTGGTGCAGAGAGAAAAAGACAGTGGGAATAGAGACTTTGTTCC	2844
Db	7758	AGGCAAAAGAGAAGTGGTGCAGAGAGAAAAAGACAGTGGGAATAGAGACTTTGTTCC	7817
QY	2845	TTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCTGCACGTCAATACGCTGACGGTAC	2904
Db	7818	TTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCTGCACGTCAATACGCTGACGGTAC	7877
QY	2905	AGGCGAGACAAATTAATGCTGATATAGGACGACGAGCAACAAATTTGCTGAGGCTATTTG	2964
Db	7878	AGGCGAGACAAATTAATGCTGATATAGGACGACGAGCAACAAATTTGCTGAGGCTATTTG	7937
QY	2965	AGGGCAACAGACATCTGTTCACATCAACAGTCTGGGGCATCAAAACAGCTCCAGGCAAGAA	3024
Db	7938	AGGGCAACAGACATCTGTTCACATCAACAGTCTGGGGCATCAAAACAGCTCCAGGCAAGAA	7997

QY 3025 TCCTGGCTGTGGAAAGATNACTTAAGATCAACAGCTCTGGGATTTGGGCTGCTCTG 3084
|||||
Db 7998 TCCGCGCTGTGGAAAGATNACTTAAGATCAACAGCTCTGGGATTTGGGCTGCTCTG 8057
QY 3085 GAAACTCATTTGGACCACTGCTGTGCTTGAATGCTAGTTGGATTAATTAATCTCGG 3144
|||||
Db 8058 GAAACATCATTTGGACCACTGCTGTGCTTGAATGCTAGTTGGATTAATTAATCTCTGG 8117
QY 3145 AACAGATTTGGAAATPAACATGACCTGGATGGAGTGGGACAGAAATTAACATTAACACAA 3204
|||||
Db 8118 AACAGATTTGGAAATPAACATGACCTGGATGGAGTGGGACAGAAATTAACATTAACACAA 8177
QY 3205 GCTTAATACACTCTCTTAATTAAGAAATCGCAAAACACGACAAAGAAATGAACAGAT 3264
|||||
Db 8178 GCTTAATACACTCTCTTAATTAAGAAATCGCAAAACACGACAAAGAAATGAACAGAT 8237
QY 3265 TATTGAATTTAGATTAATGAGCAAGTTTGTGAATTTGTGAATTTACATTAACATTTGGCTGT 3324
|||||
Db 8238 TATTGAATTTAGATTAATGAGCAAGTTTGTGAATTTGTGAATTTACATTAACATTTGGCTGT 8297
QY 3325 GGTATATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTG 3384
|||||
Db 8298 GGTATATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTG 8357
QY 3385 CTGTAATTTCTATAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTG 3444
|||||
Db 8358 CTGTAATTTCTATAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTG 8417
QY 3445 ACCTCCCAATCCCGAGGAGGAGCCGACAGAGCCCGAAGATTAAGAAAGAGTGGAGAGA 3504
|||||
Db 8418 ACCTCCCAATCCCGAGGAGGAGCCGACAGAGCCCGAAGATTAAGAAAGAGTGGAGAGA 8477
QY 3505 GAGACAGAGACAGATTCATTCGATTAAGTGAACGATTCCTTAAGCACTTATCGGAGCATC 3564
|||||
Db 8478 GAGACAGAGACAGATTCATTCGATTAAGTGAACGATTCCTTAAGCACTTATCGGAGCATC 8537
QY 3565 TGGGAGGCTGTGCTCTCTGACGTAACACCGCTTGAGAGACTTACTCTGATTTGAACGA 3624
|||||
Db 8538 TGGGAGGCTGTGCTCTCTGACGTAACACCGCTTGAGAGACTTACTCTGATTTGAACGA 8597
QY 3625 GGATTTGGAATCTCTGGAGACGAGGGGTTGGGAAGCCCTCAATATTTGGTGAATCTCC 3684
|||||
Db 8598 GGATTTGGAATCTCTGGAGACGAGGGGTTGGGAAGCCCTCAATATTTGGTGAATCTCC 8657
QY 3685 TACAGTATTTGAGTCTAGAGACTTAAGAAATAGTCTGTTAACTTCTCTCAATGCCACAGCCA 3744
|||||
Db 8658 TACAGTATTTGAGTCTAGAGACTTAAGAAATAGTCTGTTAACTTCTCTCAATGCCACAGCCA 8717
QY 3745 TAGCAGTAGCTGAGGGGACAGATAGGCTTATAGAAATTTACAAGCAAGCTTATAGAGCTA 3804
|||||
Db 8718 TAGCAGTAGCTGAGGGGACAGATAGGCTTATAGAAATTTACAAGCAAGCTTATAGAGCTA 8777
QY 3805 TTCCGCCACATACCTAGAAATTAAGACAGGGCTTGGAAGGATTTTGTATTAAGTGGGT 3864
|||||
Db 8778 TTCCGCCACATACCTAGAAATTAAGACAGGGCTTGGAAGGATTTTGTATTAAGTGGGT 8837
QY 3865 GGCAGATGCTCAAAAAGTAGTGTGATTTGATGGCTGCTGTTAAGGAAAAGATGAGACGA 3924
|||||
Db 8838 GGCAGATGCTCAAAAAGTAGTGTGATTTGATGGCTGCTGTTAAGGAAAAGATGAGACGA 8897
QY 3925 GCTGAGCCAGCAGCAGATGGGGTGGGAGCACTATCTCAGATCTAGA 3971
|||||
Db 8898 GCTGAGCCAGCAGCAGATGGGGTGGGAGCACTATCTCAGACCTAGA 8944

RESULT 9
AF070521
LOCUS AF070521 9699 bp DNA linear VRL 04-AUG-1998
DEFINITION HIV-1 E9 from the USA, complete genome.
ACCESSION AF070521
VERSION AF070521.1 GI:3378121
KEYWORDS

SOURCE
ORGANISM Human immunodeficiency virus 1 (HIV-1)
REFERENCE
AUTHORS Viruses; Retroviruses; Retroviridae; Lentivirus; Primate
TITLE lenthivirus group.
JOURNAL
MEDLINE 1 (bases 1 to 9699)
PUBMED Fang, H. and Pincus, S.H.
AUTHORS Unique insertion sequence and pattern of CD4 expression in variants
TITLE selected with immunotoxins from human immunodeficiency virus type
1-infected T cells
JOURNAL J. Virol. 69 (1), 75-81 (1995)
MEDLINE 95074930
PUBMED 7983770
AUTHORS 2 (bases 1 to 9699)
TITLE Fang, H. and Pincus, S.H.
REFERENCE Spontaneous activation of human immunodeficiency virus type 1 in an
AUTHORS Immunotoxin-resistant variant T cell line
TITLE Unpublished
JOURNAL 3 (bases 1 to 9699)
AUTHORS Fang, H. and Pincus, S.H.
REFERENCE Direct Submission
AUTHORS Submitted (04-JUN-1998) Microbiology, Montana State University, 109
JOURNAL Lewis Hall (MSU), Bozeman, MT 59717, USA
FEATURES
source
1..9699
/organism="Human immunodeficiency virus 1"
/proviral
/mol_type="genomic DNA"
/db_xref="taxon:11676"
/country="USA"
/note="Proximal sequence from the cell line E9; E9 is a
variant cell line selected by anti-HIV gp160 immunotoxin,
CD4-PE40 treatment of the cell line H9/NI4-3"
790..2292
/gene="gag"
790..2292
/gene="gag"
/codon_start=1
/product="gag protein"
/protein_id="AAC28445.1"
/db_xref="GI:3378122"
/translation="MGARASVLSGGELDKMEKIRLPQGRKLYLKHTIIMASRELER
AVNGLLETSEGRQILGLOLPSLOTGSEELRSYNTTAALYCHORIDVOTKEALD
KIEEONKSKKAOAAADTGNMSONVSPITVONLGOAMHMKETINPEAAWMDLHPV
EKASPEVTPMPSALSGBATRODINTMTNTHGAPPAIMPTIPVCEIKRTITLCKIVRA
HAGPIAPQMRPEKPSDITAGTITLQEDIGMTNHPPTIPVCEIKRTITLCKIVRA
VSPSTILDIRQPKPEPRDYDREYKTLRAQASQEVKNNMTETLLVQNPADDETIL
KALDPGATLEEMTACQGVGPGHARVLAQMSQVNPATIMTOKGNFRQRTVKC
FNCGEHIAKNCRAPRKCGMKCGECHOXKDETEROANFLGRIMPSHKRPGNFQ
SRPPTAPRESFRFEETTPPSQKPEIDKELVPLASRLSFLSPSPSQ"
<2085..5096
/gene="pol"
<2085..5096
/gene="pol"
/codon_start=1
/product="pol protein"
/protein_id="AAC28446.1"
/db_xref="GI:3378123"
/translation="FFREDLAFPOGKAREFSESQTRANSPTRRELQVWGRDNLSLEA
GADRGVSEFSPQITLMORPLVTKIGGQLEKALDGLGADTVLEENMLGRRPKM
IGGIGFVKVROYDQILIEICGHAIGVILGPPVNIIGNLTLQICSTLNPISPI
ETVPYKLRPGVDGPKVQWMPLEEKIKALVICTHMEKEGKISKTIGPNPNYTPFAI
KKKSTKRKLVDRELKRTODRHEVQGLTPHAGLKOKSVYLVVDGDAVFSVPLD
KDFKRYATFTIPINNPPGIRYOYNVLPQWKSPPALFQCSMTKILEPKNQNDIV
IYOYMDLVYGSDELEIGHRRKIELRQHLRWGFTIPDKRHOKEPPLIMGYELVPL
KMTYQPIVLPEKDSWTVNDIQLVKGLMMAISOIYAGIROLCKLIRGRTKALTEVPL
TEBALELAENREILKEPVHVVYDPSDKLAEIOKOGQWOTYIOEPRNLKTGK
YARMKGAHTNDVQKLTETAVOKIATESIYIMGKTPEKPILOKETMEAWMPYQVATVI
PEWEPVNPPLVYKLTQLEKEPIIGAEFTFYDGAANRETKSGAKGYTVDRGRQVPL
TDFTNQKTELOAHILALQDSSSEVNIYVQYALGIIQAQPKSESSELVSQITIDQLE
KEKYLLAWPAHKGIGENQOYDGLVSAQIRKVLPLDGLDKAQEBHEKHSMMRAMSD
FNLPPVNAKEIYASCDKQLEKAMHGOVDCSPGIMQDICTHLEKGLVLLVAVHAGS
IEAEVYPAETGQETAYFLKTLAGRPVYATVDNGSNFTSTTVKAAACWAGIKQEFGI
PYNQSGVIESMKNELKLTIGYVRDQAEHLKTVAHNAVFTIHNKRKKGIGIGSNGER

QY	1951	GTGTTGGATCTCTAAATNTGTAATTAAGNCGTTCATGGAAACAGACGACATGTCAATG	2010
Dp	6869	GTGTTGGATCTCTAAATNTGTAATTAAGAGGTTCAATGGAACAGACGACATGTCAATG	6928
QY	2011	TCACGACAGTTCGAATGTACACATGGAATTCAGGCGCAGTAGTCAACTCAACTGCTGTAA	2070
Dp	6929	TCACGACAGTTCGAATGTACACATGGAATTCAGGCGCAGTAGTCAACTCAACTGCTGTAA	6988
QY	2071	ATGGCAGCTTAGCAGAAAGAGATGTAGTAATTTAGATCTGCCAATTTGACAGACATGCTA	2130
Dp	6989	ATGGCAGCTTAGCAGAGAAAGATGTAGTAATTTAGATCTGCCAATTTGACAGACATGCTA	7048
QY	2131	AAACCAATNTGTACACCTGAAACACACTGTAGAAATTAATNTGTACAAAGCCCAACA	2190
Dp	7049	AAACCAATNTGTACACCTGAAACACACTGTAGAAATTAATNTGTACAAAGCCCAACA	7108
QY	2191	ATACAAAGAAAAGATCCGATCCAGAGGGGAGCCAGGAGAGCAATTTGTTACATAGAA	2250
Dp	7109	ATACAAAGAAAAGATCCGATCCAGAGGGGAGCCAGGAGAGCAATTTGTTACATAGAA	7168
QY	2251	AAATAGGAATATGAGACAAGCACATTTGTACATTTAGAGCAAAATGGAATGCCACTT	2310
Dp	7169	AAATAGGAATATGAGACAAGCACATTTGTACATTTAGAGCAAAATGGAATGCCACTT	7228
QY	2311	TAAACAGATAGCTACGCAAAATTAAGAGACACATTTGGAATATATAAAACAATATCTTTA	2370
Dp	7229	TAAACAGATAGCTACGCAAAATTAAGAGACACATTTGGAATATATAAAACAATATCTTTA	7288
QY	2371	AGCAATCCTCAGAGGGGGACCAGAAAATTTGACGACAGTTTAAATTTGGAGGGGAA	2430
Dp	7289	AGCAATCCTCAGAGGGGGACCAGAAAATTTGACGACAGTTTAAATTTGGAGGGGAA	7348
QY	2431	TTTTCTACTGTAAATTCACAACACACTGTTTAATAGTACTTGGTTTAATAGTACTTGAGTA	2490
Dp	7349	TTTTCTACTGTAAATTCACAACACACTGTTTAATAGTACTTGGTTTAATAGTACTTGAGTA	7408
QY	2491	CTGAGAGGTCAAATAACACTGGAAGAGTGTACACAACTCACACTCCCATGCAAGATTAAC	2550
Dp	7409	CTGAGAGGTCAAATAACACTGGAAGAGTGTACACAACTCACACTCCCATGCAAGATTAAC	7468
QY	2551	AATTTATTAACATGTGCGAGAAATGAGAAAAGCAATGTATGCCCCCTCCATCGTGGAC	2610
Dp	7469	AATTTATTAACATGTGCGAGAAATGAGAAAAGCAATGTATGCCCCCTCCATCGTGGAC	7528
QY	2611	AAATTTAGATGTTCACTCAAAATTTACTGGCTCTATTAAACAAGATGTGTGTAATACA	2670
Dp	7529	AAATTTAGATGTTCACTCAAAATTTACTGGCTCTATTAAACAAGATGTGTGTAATACA	7588
QY	2671	ACAATGGCTCCGAGATCTTCAGACCTGAGAGGCGATATGAGGGACAAATTGAGAAATG	2730
Dp	7589	ACAATGGCTCCCAATCTTCAGACCTGAGAGGCGATATGAGGGACAAATTGAGAAATG	7648
QY	2731	AATTTATTTAAATATTAAGTAGTAATAAATTTGACCATTAAGAGTAGACCCACCAAGGCA	2790
Dp	7649	AATTTATTTAAATATTAAGTAGTAATAAATTTGACCATTAAGAGTAGACCCACCAAGGCA	7708
QY	2791	AGAGAAAGTGTGTACAGAGAAAAAAGAGAGTGGGAATGTGGACTTTGTTCTTGGGT	2850
Dp	7709	AGAGAAAGTGTGTACAGAGAAAAAAGAGAGTGGGAATGTGGACTTTGTTCTTGGGT	7768
QY	2851	TCCTTGGAGCAGCAGGAACACTATGAGGCTGACAGTCAATGACGCTGACGGTACAGGCCA	2910
Dp	7769	TCCTTGGAGCAGCAGGAACACTATGAGGCGGAGTCAATGACGCTGACGGTACAGGCCA	7828
QY	2911	GACAAATATTTGTCATGATATAGTGTGAGACGACGAACAAATTTGCTAGAGGCTATTAAGCGC	2970
Dp	7829	GACAAATATTTGTCATGATATAGTGTGAGACGACGAACAAATTTGCTAGAGGCTATTAAGCGC	7888
QY	2971	AACAGCAATGTTTCAACTCACTGTGGGCAATCAAAACAGCTCCAGGCAAGATCTCGG	3030
Dp	7889	AACAGCAATCTTTCACACTCACTGTGGGCAATCAAAACAGCTCCAGGCAAGATCTCGG	7948
QY	3031	CTGTGGAAGATACCTTAAGAGATACACAGCTCTGGGGAATTTGGGTTGCTCTGGAAC	3090

[illegible]

Db 7181 ATACAACAATGAGTCCGAGATCTTCAGACCTGGAGGAGAGATATGAGACAATTTGA 7240
Oy 2725 GAATGGAATTTATTAATAATATAAGTACGAAAAATTTGAACCATTTAGAGTACGACCCACCA 2784
Db 7241 GAAATGAATTTATTAATAATATAAGTACGAAAAATTTGAACCATTTAGAGTACGACCCACCA 7300
Oy 2785 AGGCAAGAGAGAAGTGGTGGCAGAGAGAAAAAGAGCGTGGGAATAGAGCTTTGTTC 2844
Db 7301 AGGCAAGAGAGAAGTGGTGGCAGAGAGAAAAAGAGCGTGGGAATAGAGCTTTGTTC 7360
Oy 2845 TTGGGTTCTTGGGAGCAGACAGAGAACACTATGGGCTGCACGTCAATGACGCTGACGCTAC 2904
Db 7361 TTGGGTTCTTGGGAGCAGACAGAGAACACTATGGGCGACGCTCAATGACGCTGACGCTAC 7420
Oy 2905 AGGCGACACATTTATTTCTCTATATAGTGCACGACGACAGACAATTTGCTGAGGCTATTG 2964
Db 7421 AGGCGACACATTTATTTCTCTATATAGTGCACGACGACAGACAATTTGCTGAGGCTATTG 7480
Oy 2965 AGGCGACACACATCTGTTGCACGCTGCAGGCGATCAACAGCTCCAGGCGAAGAA 3024
Db 7481 AGGCGACACACATCTGTTGCACGCTGCAGGCGATCAACAGGCTCCAGGCGAAGAA 7540
Oy 3025 TCTGCGCTGTGAGAAAGATATCAAGATCAACAGCTCTGGGATTTGGGGTTGCTCTG 3084
Db 7541 TCTGCGCTGTGAGAAAGATATCAAGATCAACAGCTCTGGGATTTGGGGTTGCTCTG 7600
Oy 3085 GAAACATCATTTGCACACCTGCTGCTTGGATGCTAGTTGAGATATTAATCTCTGG 3144
Db 7601 GAAACATCATTTGCACACCTGCTGCTTGGATGCTAGTTGAGATATTAATCTCTGG 7660
Oy 3145 AACGATTTGGATTAACATGACGCTGATGAGTGGGAGAGAGAAATTTACAATTAACAA 3204
Db 7661 AACGATTTGGATTAACATGACGCTGATGAGTGGGAGAGAGAAATTTACAATTAACAA 7720
Oy 3205 GCTTAATACACTCTCTTAATTAAGATTCGCAAAACCAACAGAAAGATGACAGAAAT 3264
Db 7721 GCTTAATACACTCTCTTAATTAAGATTCGCAAAACCAACAGAAAGATGACAGAAAT 7780
Oy 3265 TATTGGAATTAATTAATGCGCAAGTTTGGAAATTTGTTTAACATTAACAAATTTGCTGT 3324
Db 7781 TATTGGAATTAATTAATGCGCAAGTTTGGAAATTTGTTTAACATTAACAAATTTGCTGT 7840
Oy 3325 GGTATATTAATTTATTCATATAGTATAGAGAGGCTGTGATTAAGATTAAGTTTGG 3384
Db 7841 GGTATATTAATTTATTCATATAGTATAGAGAGGCTGTGATTAAGATTAAGTTTGG 7900
Oy 3385 CTGACTTTCTATAGTATAGTATAGAGAGGATTTACACATTAATGCTTTGACAGCC 3444
Db 7901 CTGACTTTCTATAGTATAGTATAGAGAGGATTTACACATTAATGCTTTGACAGCC 7960
Oy 3445 ACCTCCCAATCCGAGGAGGAGCCGACAGCCGAGAGAAATAGAAAGAGAGTGGAGAGA 3504
Db 7961 ACCTCCCAATCCGAGGAGGAGCCGACAGCCGAGAGAAATAGAAAGAGAGTGGAGAGA 8020
Oy 3505 GAGACAGAGACAGATCCATTTGATTAAGTGAACGGATCTCTTACACTTTCTGGGAGCTAT 3564
Db 8021 GAGACAGAGACAGATCCATTTGATTAAGTGAACGGATCTCTTACACTTTCTGGGAGCTAT 8080
Oy 3565 TGGCGAGGCTGTGCTCTTACAGTACACCGCTTGAGAGACTTACTTTGATTTGAAGA 3624
Db 8081 TGGCGAGGCTGTGCTCTTACAGTACACCGCTTGAGAGACTTACTTTGATTTGAAGA 8140
Oy 3625 GGATTTGTGAACCTTGGGAGCGCAGGGGGTGGAAAGCCCTCAATATTTGGTGAATCTCC 3684
Db 8141 GGATTTGTGAACCTTGGGAGCGCAGGGGGTGGAAAGCCCTCAATATTTGGTGAATCTCC 8200
Oy 3685 TACATATTTGAGTACAGAACTTAAGAATAGTGGCTTAACCTGCTCAATCCACAGCA 3744
Db 8201 TACATATTTGAGTACAGAACTTAAGAATAGTGGCTTAACCTGCTCAATCCACAGCA 8260
Oy 3745 TAGCAGTAGCTGAGGGGACAGATAGGCTTATAGAAGTATTAACAGCAGCTTATAGAGCTA 3804
Db 8261 TAGCAGTAGCTGAGGGGACAGATAGGCTTATAGAAGTATTAACAGCAGCTTATAGAGCTA 8320

Oy 3805 TTCGCCACATACCTAGAGAATTAAGACAGAGGCTTGGAAAGATTTTGTATTAAGTGGGT 3864
Db 8321 TTCGCCACATACCTAGAGAATTAAGACAGAGGCTTGGAAAGATTTTGTATTAAGTGGGT 8380
Oy 3865 GCGAAGTGGTCAAAAAGTATGTTGATTTGATGGGCTGCTGTAAGGAAAGATGAGACGA 3924
Db 8381 GCGAAGTGGTCAAAAAGTATGTTGATTTGATGGGCTGCTGTAAGGAAAGATGAGACGA 8440
Oy 3925 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGTATCTCGAGATCTAGA 3971
Db 8441 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGTATCTCGAGCCTPAGA 8487
RESULT 11
104549 104549 9213 bp DNA linear PAT 02-DEC-1994
LOCUS Sequence 11 from Patent EP 0187041.
DEFINITION 104549
ACCESSION 104549
VERSION 104549.1 GI:591641
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 9213)
AUTHORS
Capon,D.J. and Lasky,L.A.
TITLE
Molecularly cloned acquired immunodeficiency syndrome polypeptides,
intermediates therefor and methods and materials for their use
JOURNAL
Patent: EP 0187041-A1 11 09-JUL-1986;
FEATURES
source location/Qualifiers
1..9213
BASE COUNT 3297 a 1655 c 2217 g 2044 t
ORIGIN
Query Match 41.7%; Score 2599.4; DB 6; Length 9213;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 41; Indels 6; Gaps 1;
Oy 1291 CCATGAGAGTGAAGGAGATATCAGCATCTTGTGGAGATGGGGTGAATGGGCGACCA 1350
Db 5802 CAATGAGAGTGAAGGAGATATCAGCATCTTGTGGAGATGGGGTGAATGGGCGACCA 5861
Oy 1351 TGCCTCTGGGATATGATGATCTGTAGTCTACAGAAATTTGGGCTCACCGTCTATT 1410
Db 5862 TGCCTCTGGGATATGATGATCTGTAGTCTACAGAAATTTGGGCTCACCGTCTATT 5921
Oy 1411 ATGGGCTACCTGTGGAGAGAGCAACACCACTCTATTTTGTGCATCAGATGCTTAAG 1470
Db 5922 ATGGGCTACCTGTGGAGAGAGCAACACCACTCTATTTTGTGCATCAGATGCTTAAG 5981
Oy 1471 CATATGATACAGAGTACATATGTTTGGGCGACACATGCTGTGTACCCACAGCCCA 1530
Db 5982 CATATGATACAGAGTACATATGTTTGGGCGACACATGCTGTGTACCCACAGCCCA 6041
Oy 1531 ACCCACAAGATGATATTTGTAATGTGACAGAAATTTTAACTATGTGAAAAATGCA 1590
Db 6042 ACCCACAAGATGATATTTGTAATGTGACAGAAATTTTAACTATGTGAAAAATGCA 6101
Oy 1591 TGTGTAAGCAGATGATGAGATTAATACAGTTTATGGATCAAAAGCTTAAGCATGTG 1650
Db 6102 TGTGTAAGCAGATGATGAGATTAATACAGTTTATGGATCAAAAGCTTAAGCATGTG 6161
Oy 1651 TAAATTAACCCCACTGTGTAGTTTAAAGTGCACTGATTTGAAGATGATACTAATA 1710
Db 6162 TAAATTAACCCCACTGTGTAGTTTAAAGTGCACTGATTTGAAGATGATACTAATA 6221
Oy 1711 CCATATGATAGCGGAGATGATTAATGAGAAAGAGATTAATAAAATGCTCTTCA 1770
Db 6222 CCATATGATAGCGGAGATGATTAATGAGAAAGAGATTAATAAAATGCTCTTCA 6281
Oy 1771 ATATCAGCACAAGCATAGAGATTAAGTGCGAAGAAAGATATGCAATTTCTTTAAACTTG 1830
Db 1830 ATATCAGCACAAGCATAGAGATTAAGTGCGAAGAAAGATATGCAATTTCTTTAAACTTG 1830

Dh 6282 ATATACGACAGCATTAAGGTAGGTGAGAAAATAATGCAATTTTTTATTAACCTG 6341
Qy 1831 ATATAGTACCAATAGATTA-----TACAGCTTAGTGTGATTAAGTTGTAACACCTGAG 1884
Dh 6342 ATATATATACCAATAGATTAATGATTAATCTACCACTTATAGCTTGACAAAGTTGTAACCTGAG 6401
Qy 1885 TCATTAACAGAGCCTGTCAAAGAGTATCTTTGAGCCAAATTCCTATCATTTATTTGCCCC 1944
Dh 6402 TCATTAACAGAGCCTGTCAAAGAGTATCTTTGAGCCAAATTCCTATCATTTATTTGCCCC 6461
Qy 1945 CGGCTGGTTTTGGCATCTTAAATGTATTAATTAAGCTTCAATGAAGAAAGGACCATGTA 2004
Dh 6462 CGGCTGGTTTTGGCATCTTAAATGTATTAATTAAGCTTCAATGAAGAAAGGACCATGTA 6521
Qy 2005 CAATGTGACGACAGTACAAATGTACACATGGAATCAGGCCAGTAGTATCAACTCACTGTC 2064
Dh 6522 CAATGTGACGACAGTACAAATGTACACATGGAATTAAGGCCAGTAGTATCAACTCACTGTC 6581
Qy 2065 TGTTAATTTGGCAGTCTAGCAGAGAGAGATGTAATTAATTTAGTCTGCCAATTTTCACAGACA 2124
Dh 6582 TGTTAATTTGGCAGTCTAGCAGAGAGAGATGTAATTAATTTAGTCTGCCAATTTTCACAGACA 6641
Qy 2125 ATGCTAAACCATTAATGTACAGTGAACATCTGTAGAATTAATTTGTAACAGACCCA 2184
Dh 6642 ATGCTAAACCATTAATGTACAGTGAACATCTGTAGAATTAATTTGTAACAGACCCA 6701
Qy 2185 ACACAATATAGAGAAAAAGTATCCGTATCCAGAGGGGACCCAGGAGAGCATTTGTTACAA 2244
Dh 6702 ACACAATATAGAGAAAAAGTATCCGTATCCAGAGGGGACCCAGGAGAGCATTTGTTACAA 6761
Qy 2245 TAGAGAAATATAGAAATATGAGACAGACATTTGTAACATTTAGTAGAGCAAAATGGAATG 2304
Dh 6762 TAGAGAAATATAGAAATATGAGACAGACATTTGTAACATTTAGTAGAGCAAAATGGAATG 6821
Qy 2305 CCACCTTAAAAACAGATGCTGCTAATTAAGAGAAACATTTGGAATTAATTAACAAATTA 2364
Dh 6822 ACACCTTAAAAACAGATGCTGCTAATTAAGAGAAACATTTGGAATTAATTAACAAATTA 6881
Qy 2365 TCCTTAAGCAATCTCTCAGAGAGGGGACCCAGAAATTTGTAACGACAGTTTAAATTTGAGAG 2424
Dh 6882 TCCTTAAGCAATCTCTCAGAGAGGGGACCCAGAAATTTGTAACGACAGTTTAAATTTGAGAG 6941
Qy 2425 GGGAAATTTTCTACTGTAATTTCAACACAACTGTTTAATAGTACTTGGTTTAAATGTAAT 2484
Dh 6942 GGGAAATTTTCTACTGTAATTTCAACACAACTGTTTAATAGTACTTGGTTTAAATGTAAT 7001
Qy 2485 GGAATGACTGAGGGTCAAAATTAACACTGAGAGAGTGAACACATACACTCCCATGCAAGAA 2544
Dh 7002 GGAATGACTGAGGGTCAAAATTAACACTGAGAGAGTGAACACATACACTCCCATGCAAGAA 7061
Qy 2545 TAAACAATTTTATTAACATGTGGCAGAGAGTGAAGAAAAGCAATGTATGCCCTCCCATCA 2604
Dh 7062 TAAACAATTTTATTAACATGTGGCAGAGAGTGAAGAAAAGCAATGTATGCCCTCCCATCA 7121
Qy 2605 GTGACAAATTTAGATGTTTCATCAAAATTTACTGGGCTGCTATTAAACAAGAGATGTTGTA 2664
Dh 7122 GCGGCAAAATTTAGATGTTTCATCAAAATTTACTAGGGGCTGCTATTAAACAAGAGATGTTGTA 7181
Qy 2665 ATTAACAATTTGGTCCGAGATCTTCAGACCTGAGAGAGCGGATATAGAGCAATTTGGA 2724
Dh 7182 ATTAACAATTTGGTCCGAGATCTTCAGACCTGAGAGAGCGGATATAGAGCAATTTGGA 7241
Qy 2725 GAAATGAAATTTATTAATTAATTAAGTATTAATAAATTTGAACCAATTTGAGAGTGAACCCAGCA 2784
Dh 7242 GAAATGAAATTTATTAATTAATTAAGTATTAATAAATTTGAACCAATTTGAGAGTGAACCCAGCA 7301
Qy 2785 AGGCAAAAGAGAGAGTGGTCCAGAGAGAAAAAGAGAGAGTGGAAATGAGAGCTTTGTTCC 2844
Dh 7302 AGGCAAAAGAGAGAGTGGTCCAGAGAGAAAAAGAGAGAGTGGAAATGAGAGCTTTGTTCC 7361
Qy 2845 TTGGGTTCTTGGAGCAGCAGAGAAAGCACTATGGCTCAGCTCAATGACGCTGACGGTAC 2904
Dh 7362 TTGGGTTCTTGGAGCAGCAGAGAAAGCACTATGGCTCAGCTCAATGACGCTGACGGTAC 7421

Qy 2905 AGGCAACAAATTTATGTTGATTAATGAGCAGCAGAAACAATTTGCTGAGGGCTATTG 2964
Dh 7422 AGGCAACAAATTTATGTTGATTAATGAGCAGCAGAAACAATTTGCTGAGGGCTATTG 7481
Qy 2965 AGGCGCAACAGCATCTGTTGCAACTACAGCTGTGGGCTATCAAAACAGCTCCAGGCAAGAA 3024
Dh 7482 AGGCGCAACAGCATCTGTTGCAACTACAGCTGTGGGCTATCAAAACAGCTCCAGGCAAGAA 7541
Qy 3025 TCCGAGCTGTGGAAAGATTAAGATCAACAGCTGTGGGCTATTTGGGTTGCTGCTG 3084
Dh 7542 TCCGAGCTGTGGAAAGATTAAGATCAACAGCTGTGGGCTATTTGGGTTGCTGCTG 7601
Qy 3085 GAAACATCATTTGACACCACTGCTGTGCTTGGAAATGCTAGTTGGAGTAAATTAATCTGCG 3144
Dh 7602 GAAACATCATTTGACACCACTGCTGTGCTTGGAAATGCTAGTTGGAGTAAATTAATCTGCG 7661
Qy 3145 AACGATTTTGGAAATTAACATGACCTGGATGGAGTGGGACAGAGAAATTAACAAATTAACAA 3204
Dh 7662 AACGATTTTGGAAATTAACATGACCTGGATGGAGTGGGACAGAGAAATTAACAAATTAACAA 7721
Qy 3205 GCTTAATACACTCTTAAATTTGAAGAATCCGAAAACAGCAAGAAAGAAATGAACAAAGAT 3264
Dh 7722 GCTTAATACACTCTTAAATTTGAAGAATTCGAAAACAGCAAGAAAGAAATGAACAAAGAT 7781
Qy 3265 TATTTGAATTAAGATTAATTTGGCAAGTTTGTGAATTTGTTAAACATTAACAATTTGCTGT 3324
Dh 7782 TATTTGAATTAAGATTAATTTGGCAAGTTTGTGAATTTGTTAAACATTAACAATTTGCTGT 7841
Qy 3325 GGTATTAATAATTTTCAATTAATGATTAAGAGAGCTTGTAAGTTTAAGAAATAGTTTTCG 3384
Dh 7842 GGTATTAATAATTTTCAATTAATGATTAAGAGAGCTTGTAAGTTTAAGAAATAGTTTTCG 7901
Qy 3385 CTGTACTTTCTATGTAATTAAGATTAAGAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAG 3444
Dh 7902 CTGTACTTTCTATGTAATTAAGATTAAGAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAG 7961
Qy 3445 ACCTCCCAATCCGAGGGGACCCGACAGGCCCGGAAGAAATGAAGAAAGTGGAGAGA 3504
Dh 7962 ACCTCCCAATCCGAGGGGACCCGACAGGCCCGGAAGAAATGAAGAAAGTGGAGAGA 8021
Qy 3505 GAGACAGAGACAGATTCATTCATGATTAAGAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAG 3564
Dh 8022 GAGACAGAGACAGATTCATTCATGATTAAGAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAG 8081
Qy 3565 TGCAGAGCTGTGCTCTTCAAGTACACCGCTTGAGAGACTTACTCTTATTTGTAACGA 3624
Dh 8082 TGCAGAGCTGTGCTCTTCAAGTACACCGCTTGAGAGACTTACTCTTATTTGTAACGA 8141
Qy 3625 GGATTTGTGAACCTTCTGGAGACGAGGGGTTGGAAAGCCCTCAAAATTAATTTGTAATCTCC 3684
Dh 8142 GGATTTGTGAACCTTCTGGAGACGAGGGGTTGGAAAGCCCTCAAAATTAATTTGTAATCTCC 8201
Qy 3685 TACAGTATTTGAGAGTCAAGAACTAAAGAAATAGTGTGTTAACTTGCTCAATGCCACAGCA 3744
Dh 8202 TACAGTATTTGAGAGTCAAGAACTAAAGAAATAGTGTGTTAACTTGCTCAATGCCACAGCA 8261
Qy 3745 TAGCAGTAGCAGAGGGGACAGATGAGGTTAGAGATTAACAAGCAAGCTTAATGAGCTA 3804
Dh 8262 TAGCAGTAGCAGAGGGGACAGATGAGGTTAGAGATTAACAAGCAAGCTTAATGAGCTA 8321
Qy 3805 TTGCGCCACATACCTAGAAAGATTAAGACAGAGGCTTGGAAGGATTTTCTTAAGATGGGT 3864
Dh 8322 TTGCGCCACATACCTAGAAAGATTAAGACAGAGGCTTGGAAGGATTTTCTTAAGATGGGT 8381
Qy 3865 GGCAGGTGTCAAAAAGTACTGTGATTTGATGAGCTGCTGTAAAGGAAGAAATGACAGCA 3924
Dh 8382 GGCAGGTGTCAAAAAGTACTGTGATTTGATGAGCTGCTGTAAAGGAAGAAATGACAGCA 8441
Qy 3925 GCTGAGCCACAGAGATTTGGGTTGGAGCAGTATCTGGAATCTAGA 3971
Dh 8442 GCTGAGCCACAGAGATTTGGGTTGGAGCAGTATCTGGAATCTAGA 8488

[illegible]

Db 8116 AACGATTTGGAAATACATGACCTGGATGGATGGACAGCAAAATTACAAATTACACAA 8175
 QY 3205 GCTTAATACACTCCTTAATTAGAGAAATCGCAAAACACGACAGAAAGATGACAGAAAT 3264
 Db 8176 GCTTAATACACTCCTTAATTAGAGAAATCGCAAAACACGACAGAAAGATGACAGAAAT 8235
 QY 3265 TATTGGAATTAGATTAATGGCAAGTTTGGAATTGGTTTACATAACAATTGGCTGT 3324
 Db 8236 TATTGGAATTAGATTAATGGCAAGTTTGGAATTGGTTTACATAACAATTGGCTGT 8295
 QY 3325 GGTATATAAATTTATCATATGATAGAGAGGCTGGTATAGAGTTTAAAGATAGTTTGG 3384
 Db 8296 GGTATATAAATTTATCATATGATAGAGAGGCTGGTATAGAGTTTAAAGATAGTTTGG 8355
 QY 3385 CTGTACTTTCTATAGTAATAGATTAGGAGGGATATTCACATTAATCTGTTTCAGACCC 3444
 Db 8356 CTGTACTTTCTATAGTAATAGATTAGGAGGGATATTCACATTAATCTGTTTCAGACCC 8415
 QY 3445 ACCTCCCAATCCGAGGGGAGCCGACAGGCCGGAAGAAATGAAGAAGAGTGGAGAGA 3504
 Db 8416 ACCTCCCAATCCGAGGGGAGCCGACAGGCCGGAAGAAATGAAGAAGAGTGGAGAGA 8475
 QY 3505 GAGACAGAGACAGATCCATTGATTAAGAAAGGATCTTACACTTATCTGGAGCATC 3564
 Db 8476 GAGACAGAGACAGATCCATTGATTAAGAAAGGATCTTACACTTATCTGGAGCATC 8535
 QY 3565 TGGGAGACCTGTGCTCTTTCAGCTACACCGCTTGAGAGACTTACTCTGTTTAAACA 3624
 Db 8536 TGGGAGACCTGTGCTCTTTCAGCTACACCGCTTGAGAGACTTACTCTGTTTAAACA 8595
 QY 3625 GGATGTGGAACCTCTGGAGCGGAGGGGGTGGGAAGCCCAATATATGGTGAATCTCC 3684
 Db 8596 GGATGTGGAACCTCTGGAGCGGAGGGGGTGGGAAGCCCAATATATGGTGAATCTCC 8655
 QY 3685 TACAGTATGAGAGTCAGAGAACTAAGAATAGTGTCTTAACTTGCTCAATGCCACAGCA 3744
 Db 8656 TACAGTATGAGAGTCAGAGAACTAAGAATAGTGTCTTAACTTGCTCAATGCCACAGCA 8715
 QY 3745 TAGCAGTACTGAGAGGGACAGATAGGTTATAGAATTAACAAGACCTTATAGACTA 3804
 Db 8716 TAGCAGTACTGAGAGGGACAGATAGGTTATAGAATTAACAAGACCTTATAGACTA 8775
 QY 3805 TTGCGCACATACCTAGAGAAATAGACAGGGCTTGAAGAGATTTCTCTAATAAGATGGT 3864
 Db 8776 TTGCGCACATACCTAGAGAAATAGACAGGGCTTGAAGAGATTTCTCTAATAAGATGGT 8835
 QY 3865 GGCAGATGTCCTCAAAAAGTAGTGTGATTTGATGGCTGCTGTAAGAGAAAGATGAGACA 3924
 Db 8836 GGCAGATGTCCTCAAAAAGTAGTGTGATTTGATGGCTGCTGTAAGAGAAAGATGAGACA 8895
 QY 3925 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGTATCTGAGATCTAGA 3971
 Db 8896 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGTATCTGAGATCTAGA 8942

RESULT 13
 HIVPV22 9770 bp DNA linear VRL 01-OCT-1999
 LOCUS Human immunodeficiency virus type 1, isolate PV22, complete genome
 DEFINITION (H9/HTLV-III proviral DNA).
 ACCESSION K02083
 VERSION K02083.1 GI:555008
 KEYWORDS TRK protein; acquired immune deficiency syndrome; complete genome; env protein; gag protein; long terminal repeat; pol protein; polyprotein; proviral gene; rev protein; reverse transcriptase; tat protein; transactivator.
 SOURCE Human immunodeficiency virus 1 (HIV-1)
 ORGANISM Human immunodeficiency virus 1
 Viruses: Retroviral viruses; Retroviridae; Lentivirus; Primate lentivirus group.
 REFERENCE 1 (bases 1 to 9770)
 AUTHORS Muesling,M.A., Smith,D.H., Cabradilla,C.D., Benton,C.V., Lasky,L.A.

TITLE and Capon,D.J.
 JOURNAL Nucleic acid structure and expression of the human
 MEDLINE AIDS/Lymphadenopathy retrovirus
 PUBMED Nature 313 (6002), 450-458 (1985)
 85111157
 2982104
 REFERENCE 2 (sites)
 AUTHORS van Beveren,C.P., Coffin,J. and Hughes,S.
 TITLE Appendix B: HTLV-3/LAV genome
 JOURNAL (in) Weiss,R.L., Teich,N., Varmus,H. and Coffin,J. (Eds.);
 RNA TUMOR VIRUSES, MOLECULAR BIOLOGY OF TUMOR VIRUSES, SECOND
 EDITION, 2: Supplements and Appendices: 1106-1123;
 Cold Spring Harbor Laboratory, CSH, NY (1985)
 3 (bases 2111 to 2112)
 Muesling,M.A.
 REFERENCE On Oct 4, 1994 this sequence version replaced gi:328551.
 AUTHORS unpublished
 JOURNAL
 COMMENT (1) revised sequence, personal communication.
 ((in) Weiss,R., Teich,N., Varmus,H. and Coffin,J. (Eds.);RNA Tumor
 Viruses,Molecu review; bases 1 to 9769.
 [3] revises [1],((in) Weiss,R., Teich,N., Varmus,H. and Coffin,J.
 (Eds.);RNA Tumor Viruses,Molecu).
 This sequence for a H9/HTLV-III virus was determined from one
 complete proviral clone [1]. Additionally, several cDNA clones of
 the viral RNA were sequenced for comparison with the entire
 proviral sequence. The differences between cDNA and proviral DNA
 are extensive and are listed in the Sites table as variations. The
 authors believe that the variations may be due in part to different
 strains in the H9/HTLV-III cell line, because it was established by
 infection with material from several AIDS patients. With the
 addition of g at 2111, gag cds and pol cds are very close to those
 of HXB2, BRU, and related HIV viruses.
 For details and other references pertaining to Sites and Features,
 see the HIV reference entry.
 location/Qualifiers
 1..9770
 /organism="Human immunodeficiency virus 1"
 /proviral
 /mol_type="genomic DNA"
 /db_xref="taxon:11676"
 10..643
 /note="5' LTR"
 386..395
 /bound_molecly="Sp1 III"
 397..406
 /bound_molecly="Sp1 II"
 408..417
 /bound_molecly="Sp1 I"
 453..560
 /note="R repeat 5' copy"
 464..9678
 /note="genomic mRNA"
 464..9678
 /note="tat, rev, nef subgenomic mRNA"
 464
 /note="numbered 1 in [1]"
 510
 /note="a in provirus; g in cDNA [1]"
 /replace="g"
 575
 /note="g in provirus; a in cDNA [1]"
 /replace="a"
 645..662
 /bound_molecly="Lys-tRNA primer"
 753..5822
 /note="tat, rev, nef subgenomic mRNA intron 1"
 799..2337
 /note="gag polyprotein precursor"
 /codon_start=1
 /protein_id="AAB59866.1"
 /db_xref="GI:328554"
 /translation="MGARASVLSGGELDRMEKIRLRPGKKKKYKHIIVASPELEFR
 AVNPGLETSBGCQILIGOLPISLQTSBELRSIXNTVATLYCYHQRIRIKDKREALD

```
CDS
    /partial
    /note="pol polyprotein (NH2-terminus uncertain)"
    /codon_start=-1
    /protein_id="AAB59867.1"
    /db_xref="GI:328555"
    /translation="FPEEDLAFLOGKAREPSSBOTRANSPITISEOTRANSPRIEIO
    VMGRDNSSPEAGADROGYSFNPOTTLMOBRLVTKIGGOKLEALDGDADYIE
    EMLPGRMKRMKIGGIGFIRKYOYDILEITGKRAIGVIVLGPPTVNTIGRNLTO
    IGCTLPNPIPIETVPVTKPKGMDGPVKQWPKLEEKIKALVEICTEMKEGKISKIG
    PENRYNTPVPAIKKDKSTWKRLVDFRELNRKQDEWVLOGIIPHAGLKRSVTVL
    DVDDAFSVPLDEDFKRYTAFIIPSIINNETPGIRYQYNVLPQMGKSPALFQSSMTI
    LEPFRKONPDIVITYOYMDLVGSDLEIGHRTKIEELRHLRMGLTTPDKKHOKEP
    PELMGCELDHDKWTQPIVLPERKDSWTVNDIOKNGKLMMASOIYPIGIVRDLCKL
    RGTALTEVPLTBEAELEAENREILKEPVHYGYIDPSKDLAETOKOGOGQWTOI
    YQEPFNKLTGKTAARMGATNDVKQLEAVQRTTESIVIMGKTPKFLPIQKEME
    TWMEYQWATWIMEPEFVNTPPLVKMLYOLEKEPIVGAETFYVDGAANRETRIGKAGY
    LTNKGRQKVPLNTNTNOKETLOAIYALQDSLEVINIVDSOYALGIIQADPDQSES
    ELVNOIIEOLIKROYVLAHPAKHGIGGNEOYDKLASIRKILFLDGIIDKODEHE
    KYHNSNRAMASDENLPPVYAKETIYASDCQKOLGEAMHGOVCTTDSNGWSTCHLEK
    VIVIVAVASGYIEAEYIPAEIGOTAFIPLIKTAGRPVPTTHDNGSNATSTATKRA
    CMWAGIKQERGIPTNPQSGVSESNMELKIIQVARDQEHKLTAVQMAVPIHNERK
    KGGVIGGSAEIRVDIITADIQRELOKQITKIONFVYRDSRNP.LMKGPAPLMLWG
    EGAVVIODNSDIIVPRRKAKIIRDYGKQMAAGDVCVASRODED"
    2111..2112
    /ciliation=[1]
    5086..5664
    /note="vif protein"
    /codon_start=-1
    /protein_id="AAB59868.1"
    /db_xref="GI:328556"
    /translation="MENRQVMIYMOYDRMRIRTKSLYKHHMYSGKANGWYRHHY
    ESPHRISSEVHPIGDARLVITTYWGLTHGIDRPHMGVGVSEMRKRRSTQVDPDL
    ADOLIHLYFDSCDSKIRKALLGHIVSPCEYQAGHNKVSLOYLALALITPKIK
    PLPVSYTKLEDNMNKPQKTKGRHSGHTMNGH"
    5604..5840
    /note="vpr protein"
    /codon_start=-1
    /protein_id="AAB59869.1"
    /db_xref="GI:328557"
    /translation="MEQAPEDGQPRQREPHNEWTLELLEELKNEAVRHPRIWLHGLGQ
    HIYETYGDTWAGVEAIRILQDLFIHFQMWST"
    5716
    /note="g in provirus; a in cDNA [1]"
    /replace="a"
    join(5876..6090,8421..8466)
    /codon_start=-1
    /note="tat protein"
    /codon_start=-1
    /protein_id="AAB59870.1"
    /db_xref="GI:328552"
    /translation="MEPVDRLEPMKHNPSOPKTACNCCCKCFHCQYCFITKALG
    IYGRKKRRRORRRPQGSQTHQVSLSKOPTSOPRGDPTGKE"
    5992
    /note="a in provirus; g in cDNA [1]"
    /replace="g"
    6007
    /note="c in provirus; t in cDNA [1]"
    /replace="t"
    join(6015..6090,8421..8695)
    /note="rev protein"
    /codon_start=-1
    /protein_id="AAB59871.1"
    /db_xref="GI:328553"
    /translation="MAGRSGDSDEDLKAVRLIKFLYQSNPPNPECTROARRNRRR
    WRBROHISHISERILSYLIGRSAEPVQLPLERLITLDCNEDCGTSGQVGSPOI
```

```
variation
    /note="c in provirus; g in cDNA [1]"
    /replace="g"
    6051
    /note="c in provirus; a in cDNA [1]"
    /replace="a"
    6055..6057
    /note="agg in provirus; gaa in cDNA [1]"
    /replace="gaa"
    6091..8420
    /note="tat cds intron 2"
    6091..8420
    /note="rev cds intron 2"
    6091..8420
    /note="tat, rev, nef subgenomic mRNA intron 2"
    6107..6352
    /note="vpu protein"
    /codon_start=-1
    /protein_id="AAB59872.1"
    /db_xref="GI:328558"
    /translation="MOPIDATVATVATIIATVMSYIVIEYKILRQKIDRLDR
    LIRAEQDSNESEGETSALVEKGVEMGHAPWVDL"
    6108
    /note="t in provirus; c in cDNA [1]"
    /replace="c"
    6120
    /note="a in provirus; c in cDNA [1]"
    /replace="c"
    6125..6126
    /note="gc in provirus; gtaac in cDNA [1]"
    /replace="gtaac"
    6136
    /note="a in provirus; c in cDNA [1]"
    /replace="c"
    6235
    /note="t in provirus; a in cDNA [1]"
    /replace="a"
    6267..8837
    CDS
    Query Match 41.7%; Score 2599.4; DB 14; Length 9770;
    Best Local Similarity 98.3%; Pred. No. 0;
    Matches 2640; Conservative 0; Mismatches 41; Indels 6; Gaps 1;
    QY 1291 CCATGAGAGTGAAGAGAAATATCAGACTTGTGAGATGGGGTGAATGGGCACCA 1350
    Db 6265 CAATGAGAGTGAAGAGAAATATCAGACTTGTGAGATGGGGTGAATGGGCACCA 6324
    QY 1351 TGCTCCTTGGGATATTGATGATCTGTAGTGTACAGAAAAATTTGGGTCACCGCTATT 1410
    Db 6325 TGCTCCTTGGGATATTGATGATCTGTAGTGTACAGAAAAATTTGGGTCACAGCTATT 6384
    QY 1411 ATGGGGTACTGTGTGGAAGAGCAACCACTCTATTTTGTGATGATGATGATGATGAT 1470
    Db 6385 ATGGGGTACTGTGTGGAAGAGCAACCACTCTATTTTGTGATGATGATGATGATGAT 6444
    QY 1471 CATATGATACAGAGTACATATATGTTGGCCACACATGCCCTGTATACCCAGACACCA 1530
    Db 6445 CATATGATACAGAGTACATATATGTTGGCCACACATGCCCTGTATACCCAGACACCA 6504
    QY 1531 ACCCACAAGAGTATGTTGTAATGTGACAGAAAAATTTTAAATCATGTGGAATAATGACA 1590
    Db 6505 ACCCACAAGAGTATGTTGTAATGTGACAGAAAAATTTTAAATCATGTGGAATAATGACA 6564
    QY 1591 TGGTAGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1650
    Db 6565 TGGTAGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6624
    QY 1651 TAAATTTAACCCACCTGTGTAGTTTAAAGTGCACCTGATTTGAAGATGATGATGATGAT 1710
    Db 6625 TAAATTTAACCCACCTGTGTAGTTTAAAGTGCACCTGATTTGAAGATGATGATGATGAT 6684
    QY 1711 CCAATAGTACTACGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1770
```

|||||
Db 6685 CCAATAGTAGAGGGGAGATGATTAATGGAAGAAAGAGATATAAAACGCTCTTCA 6744
Oy 1771 ATATCAGCACAAGATTAAGAGATAGTGCAGAAAAGAAATATGCAATCTTTTATAAAGCTG 1830
Db 6745 ATATCAGCACAAGATTAAGAGATAGTGCAGAAAAGAAATATGCAATCTTTTATAAAGCTG 6804
Oy 1831 ATATAGTAGCAATAGATA-----TACCAGTATAGTGTATAGTGTATAGCCTAG 1884
Db 6805 ATATATATACCAATAGATATATATCTACAGCTATAGCTGTGACAAGTGTAAACCTAG 6864
Oy 1885 TCATTCACAGAGCCTGTCCAAAGGTATCTTTGAGCCAAATCCCATATATTTGTGCC 1944
Db 6865 TCATTCACAGAGCCTGTCCAAAGGTATCTTTGAGCCAAATCCCATATATTTGTGCC 6924
Oy 1945 CGGCTGTTTGGCATCTTAATAATGTATTAATAAGCTGTAATGAAACAGAGCCATGTA 2004
Db 6925 CGGCTGTTTGGCATCTTAATAATGTATTAATAAGCTGTAATGAAACAGAGCCATGTA 6984
Oy 2005 CAAATGTCAGCAGTACATGTACATGTGAAATCAGAGCCAGTAGTATCACTCACTGC 2064
Db 6985 CAAATGTCAGCAGTACATGTACATGTGAAATTTAGCCAGTAGTATCACTCACTGC 7044
Oy 2065 TGTTAATGCGAGTCTAGCAGAGAAGATGTAGTAATAGATCTGCCAATTTCCAGACA 2124
Db 7045 TGTTAATGCGAGTCTAGCAGAGAAGATGTAGTAATAGATCTGCCAATTTCCAGACA 7104
Oy 2125 ATGCTAAACCATATATAGTACAGCTGAACACATCTGTAGAAATTAATGTACAAAGCCA 2184
Db 7105 ATGCTAAACCATATATAGTACAGCTGAACCAATCTGTAGAAATTAATGTACAAAGCCA 7164
Oy 2185 ACAACAATCAAGAAAAAGTATCCGTATCCAGAGGGGACAGGGGAGACATTTGTTACAA 2244
Db 7165 ACAACAATCAAGAAAAAGTATCCGTATCCAGAGGGGACAGGGGAGACATTTGTTACAA 7224
Oy 2245 TAGGAAAAATAGAAATATGAGACAGCAGCATTTGTACATTAGTAGAGCAAAATGGAATG 2304
Db 7225 TAGGAAAAATAGAAATATGAGACAGCAGCATTTGTACATTAGTAGAGCAAAATGGAATG 7284
Oy 2305 CCATTTTAAACAGATAGCTAGCAAAATTAAGAGAACATTTGGAATTAATAACATTA 2364
Db 7285 ACACCTTTAAACAGATAGTAAAGCAAAATTAAGAGAACATTTGGAATTAATAACATTA 7344
Oy 2365 TCTTTAAGCAATCCTCAGAGGGGAGCCAGAAATTTGTAAAGCAGATTTTAAATGTGAG 2424
Db 7345 TCTTTAAGCAATCCTCAGAGGGGAGCCAGAAATTTGTAAAGCAGATTTTAAATGTGAG 7404
Oy 2425 GGGAAATTTTCTACTGTAAATCAACACAACTGTTTAATAGTACTGTTTAAATAGTACTT 2484
Db 7405 GGGAAATTTTCTACTGTAAATCAACACAACTGTTTAATAGTACTGTTTAAATAGTACTT 7464
Oy 2485 GGAGTACTGAAGGGTCAAAATTAACACTGTAAGAGAGTGAACATACACTCCATGACAGAA 2544
Db 7465 GGAGTACTGAAGGGTCAAAATTAACACTGTAAGAGAGTGAACATACACTCCATGACAGAA 7524
Oy 2545 TAAACAATTTAATAACATGTGAGAGAGTAAAGAAAAGCAATGTATGCCCTCCATCA 2604
Db 7525 TAAACAATTTAATAACATGTGAGAGAGTAAAGAAAAGCAATGTATGCCCTCCATCA 7584
Oy 2605 GTGAGCAAAATAGATGTTCAATCAATATTTACTGGGCTGCTATTAAACAAGAGATGGTGA 2664
Db 7585 GCGGACAATATAGTATGTTCAATCAATATTTACAGGGCTGCTATTAAACAAGAGATGGTGA 7644
Oy 2665 ATAAACAATGSGTCCGAGATCTTCAGACCTGAGAGAGCGGATATGAGGGACAATTTGA 2724
Db 7645 ATAAACAATGSGTCCGAGATCTTCAGACCTGAGAGAGGAGATATGAGGGACAATTTGA 7704
Oy 2725 GAAATGAAATTAATAATTAATAATAGTATAGTAAAAATTGAACCTTTAGAGTAGACACCACA 2784
Db 7705 GAAATGAAATTAATAATTAATAATAGTATAGTAAAAATTGAACCTTTAGAGTAGACACCACA 7764
Oy 2785 AGGCAAGAGAGAGTGGTGCAGAGAGAAAAAGAGAGTGGAAATGAGAGCTTTGTTC 2844
|||||

7765 AGGCAAGAGAGAGTGGTGCAGAGAGAGAAAAAGAGAGTGGAAATAGAGCTTTGTTC 7824
Oy 2845 TTGGTTCTTGGAGACAGACAGAAAGCACTATAGGCTGCAGTCAATGACGCTGACGCTAC 2904
Db 7825 TTGGTTCTTGGAGACAGACAGAAAGCACTATAGGCGCGAGCTCAATGACGCTGACGCTAC 7884
Oy 2905 AGGCGAGCAATTTTGTCTATATAGTATAGGAGCAGCAGACAAATTTGCTGAGGGCTATTT 2964
Db 7885 AGGCGAGCAATTTTGTCTATATAGTATAGTATAGTACAGCAGCAGAAATTTGCTGAGGGCTATTT 7944
Oy 2965 AGGCGCAACAGCATCTGTTGCACTCACAGTCTGGGGCATCAAAAGCTCCAGGCAAGAA 3024
Db 7945 AGGCGCAACAGCATCTGTTGCACTCACAGTCTGGGGCATCAAAAGCTCCAGGCAAGAA 8004
Oy 3025 TCTGAGCTGTGAAAGATACCTTAAGATCAACAGCTCTGGGATTTGGGTTGCTCTG 3084
Db 8005 TCTGAGCTGTGAAAGATACCTTAAGATCAACAGCTCTGGGATTTGGGTTGCTCTG 8064
Oy 3085 GAAAACGATTTGCAACCACTGCTGCTGCTGGAATGCTAGTTGGATTAATCTCTGG 3144
Db 8065 GAAAACGATTTGCAACCACTGCTGCTGCTGGAATGCTAGTTGGATTAATTAATCTCTGG 8124
Oy 3145 AACAGATTTGGAATTAACATGACCTGATGAGAGTGGGACAGAGAAATTAACATTAACAA 3204
Db 8125 AACAGATTTGGAATTAACATGACCTGATGAGAGTGGGACAGAGAAATTAACATTAACAA 8184
Oy 3205 GCTTAATACACTCCTTAATTAAGAGATGCGAAACAGCAGAGAAAGATTAACAGAT 3264
Db 8185 GCTTAATACACTCCTTAATTAAGAGATGCGAAACAGCAGAGAAAGATTAACAGAT 8244
Oy 3265 TATTGGAATTTGATTAATGAGGCGCAAGTTGTGGAATTTGTTTAACTAACAAATTTGGCTGT 3324
Db 8245 TATTGGAATTTGATTAATGAGGCGCAAGTTGTGGAATTTGTTTAACTAACAAATTTGGCTGT 8304
Oy 3325 GGTATATAAATTAATTCATAATGATAGTAGAGAGGCTTGATAGTTTAAAGATTAATTTTG 3384
Db 8305 GGTATATAAATTAATTCATAATGATAGTAGAGAGGCTTGATAGTTTAAAGATTAATTTTG 8364
Oy 3385 CTGTACTTTCTATAGTAAATAGAGTTAGGCAAGGATATTCACCTATACGTTTCAGACCC 3444
Db 8365 CTGTACTTTCTATAGTAAATAGAGTTAGGCAAGGATATTCACCTATACGTTTCAGACCC 8424
Oy 3445 ACCTCCCAATCCCGAGGGGAGCCGACAGGCCCGAAGAAATGAAGAAAGAGTGGAGAGA 3504
Db 8425 ACCTCCCAATCCCGAGGGGAGCCGACAGGCCCGAAGAAATGAAGAAAGAGTGGAGAGA 8484
Oy 3505 GAGACAGAGACAGATCCATGATTAAGTGAACGGAATCCTTAGCACTTATCTGGAGCATC 3564
Db 8485 GAGACAGAGACAGATCCATGATTAAGTGAACGGAATCCTTAGCACTTATCTGGAGCATC 8544
Oy 3565 TGGGAGGCTGTGCTCTTCAGCTACACCGCCTTGAGAGACTTACTTGAATTTGAACGA 3624
Db 8545 TGGGAGGCTGTGCTCTTCAGCTACACCGCCTTGAGAGACTTACTTGAATTTGAACGA 8604
Oy 3625 GGATTTGGAAGCTCTGGGAGCGAGGGGGTGGGAAAGCCCTAAATATTTGGTGGAAATCC 3684
Db 8605 GGATTTGGAAGCTCTGGGAGCGAGGGGGTGGGAAAGCCCTAAATATTTGGTGGAAATCC 8664
Oy 3685 TACAGTATTTGGAGTCAGAGAACTAAGAAATATGCTGTAACTCTCCATCCACAGCA 3744
Db 8665 TACAGTATTTGGAGTCAGAGAGTAAAGAAATATGCTGTAGTCAAGAGACTTAATAGAGCTA 8724
Oy 3745 TAGCAGTAGCTGAGGGGACAGATAGGCTTATAGAGATTTTCAAGACGCTTATAGAGCTA 3804
Db 8725 TAGCAGTAGCTGAGGGGACAGATAGGCTTATAGAGATTTTCAAGAGACTTAATAGAGCTA 8784
Oy 3805 TTTGCCACATATCCATGAAGAAATAGACAGGGCTTGGAAGATTTTCTATTAAGATGGGT 3864
Db 8785 TTTGCCACATATCCATGAAGAAATAGACAGGGCTTGGAAGATTTTCTATTAAGATGGGT 8844
Oy 3865 GGCAGTGTGTCAAAAAGTAGTGTGATTTGGAATGGCTCTCTGTAAGGAGAAAGATGAGACGA 3924
Db 8845 GGCAGTGTGTCAAAAAGTAGTGTGATTTGGAATGGCTCTCTGTAAGGAGAAAGATGAGACGA 8904
|||||

QY 3925 GCTGAGCCAGCAGATGGGTGGAGCATATCTGAGATCTAGA 3971
|||||
Db 8905 GCTGAGCCAGCAGATGGGTGGAGCATCTGAGACCTAGA 8951

RESULT 14
E01088
LOCUS Nucleic acid sequence coding for HTLV envelop protein.
DEFINITION E01088
ACCESSION E01088
VERSION E01088.1 GI:2169347
KEYWORDS JP 1987012799-A/1.
SOURCE unidentified
ORGANISM unclassified.

REFERENCE
AUTHORS 1 (bases 1 to 3156)
Robaoto,M.K., Robaoto,C.G., Eragamu,P.R., Jiyooji,M.S. and
Furotsushii,I.U.
TITLE ENVELOPE PROTEIN OF VIRUS OF ACQUIRED IMMUNODEFICIENCY SYNDROME
JOURNAL Patent: JP 1987012799-A 1 21-JAN-1987;
F. HOFMANN LA ROCHE & CO AG, USA GOVERNMENT
COMMENT OS Human [Homo sapiens]
PN JP 1987012799-A/1
PD 21-JAN-1987
PF 18-APR-1986 JP 1986089830
PI 19-APR-1985 US 85 725021
PI ROBATO MITSUCHIERU KUROURU, ROBATO CHIYARUZU GARO, PI
ERAGAMU PUREMUKUMA REDEL, JIYOOJI METSUDO SHIYOU, PI FUROTSUSHII
IICHINGU UONGUSUTARU
PC C07K13/00,A61K39/21,C07H21/04,C12N1/20,C12N15/00,C12P21/00,PC
C12P21/02,
PC G01N33/569,G01N33/577,(C12N1/20,C12R1:125),(C12N1/20,C12R1:19),PC
(C12P21/02,
PC C12R1:125),(C12P21/02,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: library=genome library;
FH key Location/Qualifiers
FT mat_peptide 1..3157
FT Location/Qualifiers
1..3156
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 1095 a 536 c 764 g 761 t
ORIGIN

Query Match 41.7%; Score 2598.2; DB 6; Length 3156;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 38; Indels 6; Gaps 1;

QY 1291 CCATGAGAGTGAAGAGAGATATCAGCACTTGTGAGATGGGGGTGGAATGGGGACCA 1350
|||||
Db 476 CAATGAGAGTGAAGAGAGATATCAGCACTTGTGAGATGGGGGTGGAATGGGGACCA 535

QY 1351 TGCCTCTTGGGATATTATGATCTGTAGTGTGCTACAGAAAAATTTGGGTCACTGTAATT 1410
|||||
Db 536 TGCCTCTTGGGATATTATGATCTGTAGTGTGCTACAGAAAAATTTGGGTCACTGTAATT 595

QY 1411 ATGGGAGTACCTGTGTGAAGAGCAACCACTATTTTGGCATCGATGCTTAAG 1470
|||||
Db 596 ATGGGAGTACCTGTGTGAAGAGCAACCACTATTTTGGCATCGATGCTTAAG 655

QY 1471 CATATGATACAGAGTACATATGTTTGGGCACACATGCTGTGATCCACAGACCCA 1530
|||||
Db 656 CATATGATACAGAGTACATATGTTTGGGCACACATGCTGTGATCCACAGACCCA 715

QY 1531 ACCCACAAGAAGTAGTATTGTTAAATGTGACAGAAAAATTTTAACATGTGGAAAAATGACA 1590
|||||
Db 716 ACCCACAAGAAGTAGTATTGTTAAATGTGACAGAAAAATTTTAACATGTGGAAAAATGACA 775

QY 1591 TGGTAGAACAGATGATGAGATATTAATCAAGTTTATGGATTCAAAGCCTTAAGCCATGTC 1650
|||||
Db 776 TGGTAGAACAGATGATGAGATATTAATCAAGTTTATGGATTCAAAGCCTTAAGCCATGTC 835

QY 1651 TAAATTAACCCCACTGTGTAGTTAATTAAGGCACTGTGTTGAACAAATGATACATATA 1710
|||||
Db 836 TAAATTAACCCCACTGTGTAGTTAATTAAGGCACTGTGTTGAACAAATGATACATATA 895

QY 1711 CCAATAGTAGAGGGGAGAGATGATAATGAGAAAGAGATAAAAAAGCTGCTTTCA 1770
|||||
Db 896 CCAATAGTAGAGGGGAGAGATGATAATGAGAAAGAGATAAAAAAGCTGCTTTCA 955

QY 1771 ATATCAGACACAGATTAAGAGATTAAGTGCAGAAAGAAATTCATTTTATTAACCTTG 1830
|||||
Db 956 ATATCAGACACAGATTAAGAGATTAAGTGCAGAAAGAAATTCATTTTATTAACCTTG 1015

QY 1831 ATATAGTACCAATAGATAA-----TACAGCTATAGTTGATTAAGTTTAACACCTCAG 1884
|||||
Db 1016 ATATAGTACCAATAGATAATGATATCTACAGCTATAGTTGATTAAGTTTAACACCTCAG 1075

QY 1885 TCATTTACACAGGCGCTGCCAAGGTATCCTTGAGCCCAATTCATATCATTAATGTCGCC 1944
|||||
Db 1076 TCATTTACACAGGCGCTGCCAAGGTATCCTTGAGCCCAATTCATATCATTAATGTCGCC 1135

QY 1945 CGGCTGTTTGGGATTTCTAAATGTATTAATTAAGAGCTTCATGGAACGAGCAACATGTA 2004
|||||
Db 1136 CGGCTGTTTGGGATTTCTAAATGTATTAATTAAGAGCTTCATGGAACGAGCAACATGTA 1195

QY 2005 CAAATGTCACAGCAGTCAATGTATACATGGAATCAAGCCAGTATGATCACTCACTGTC 2064
|||||
Db 1196 CAAATGTCACAGCAGTCAATGTATACATGGAATCAAGCCAGTATGATCACTCACTGTC 1255

QY 2065 TGTTAATGCGAGTCTAGCAGAAAGAGATGTAATTTAGATCTGCCAATTTACAGACA 2124
|||||
Db 1256 TGTTAATGCGAGTCTAGCAGAAAGAGATGTAATTTAGATCTGCCAATTTACAGACA 1315

QY 2125 ATGCTAAACCATATATGTACAGTGCAGTCACTGTGAAAGAAATTAATTTGACAAAGCCA 2184
|||||
Db 1316 ATGCTAAACCATATATGTACAGTGCAGTCACTGTGAAAGAAATTAATTTGACAAAGCCA 1375

QY 2185 ACAACATTAACAGAAAAAGTATCCGTATCCAGAGGGGACCCAGAGACATTTGTTACAA 2244
|||||
Db 1376 ACAACATTAACAGAAAAAGTATCCGTATCCAGAGGGGACCCAGAGACATTTGTTACAA 1435

QY 2245 TAGGAAAAATAGGAATATAGACACACATTTGTAATTAATTAAGGCAAAATGGAATG 2304
|||||
Db 1436 TAGGAAAAATAGGAATATAGACACACATTTGTAATTAATTAAGGCAAAATGGAATG 1495

QY 2305 CCACTTTAAACAGATAGTACGCAAAATTAAGAGAACATTTGCAAAATTAATAAACAATA 2364
|||||
Db 1496 CCACTTTAAACAGATAGTACGCAAAATTAAGAGAACATTTGCAAAATTAATAAACAATA 1555

QY 2365 TCTTTAAGCAATCCTCAGAGAGGAGCCAGAAATTTGAACGACAGTTTAATTTGAGAG 2424
|||||
Db 1556 TCTTTAAGCAATCCTCAGAGAGGAGCCAGAAATTTGAACGACAGTTTAATTTGAGAG 1615

QY 2425 GGGAAATTTTCTACTGTAATTTCAACACACTGTTTAATTAATTAATTAATTAATTA 2484
|||||
Db 1616 GGGAAATTTTCTACTGTAATTTCAACACACTGTTTAATTAATTAATTAATTAATTA 1675

QY 2485 GGAATAGTGAAGGGTCAAAATTAACCTGAAGAGTGCACAAATCAGCTCCATGACGAA 2544
|||||
Db 1676 GGAATAGTGAAGGGTCAAAATTAACCTGAAGAGTGCACAAATCAGCTCCATGACGAA 1735

QY 2545 TAAACCAATTTTAATTAACATGTGCAGAGAAATAGGAAAGCAATGTATGCCCTCCATCA 2604
|||||
Db 1736 TAAACCAATTTTAATTAACATGTGCAGAGAAATAGGAAAGCAATGTATGCCCTCCATCA 1795

QY 2605 GTGACAAATTTAGATGTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2664

[illegible]

Hostomsky, Z., Hostomskaya, Z., Hudson, G.O., Moonaw, E.W. and Nodes, B.R. Reconstitution in vitro of RNase H activity by using purified N-terminal and C-terminal domains of human immunodeficiency virus type 1 reverse transcriptase
Proc. Natl. Acad. Sci. U.S.A. 88 (4), 1148-1152 (1991)

9144129

1705027

Original source text: Human immunodeficiency virus type 1 (HIV-1) proviral DNA clone BH10.
(1 in Weiss, R., Teich, R., Varmus, H. and Coffin, J. (Eds.): RNA Tumor Viruses, Molecu review; bases 1 to 8932.

The BH10 sequence differs from BH8 and BH5 by 0.9% in the coding regions and 1.8% in the noncoding regions, and the authors of [1] believe that these are stable variants.

The HTLV-III genome encodes at least seven proteins: gag, pol, env, tat, tax, 27k antigen and the sor 23k product. The 3' ORF (positions 8153-8773) is truncated in BH10 (stop codon at positions 8522-8524), but reads through in BH8 and other sequences to yield what is now called the 27k antigen.

The mechanism for pol gene translation has not been elucidated: a gag-pol fusion protein is possible; splicing or frameshift have not been ruled out. The viral protease would be determined by the

region in question.
 The Tat protein (trans-activator protein, approximately 14 kb) is an effector of an autostimulatory pathway through interaction with a positive control element, the trans-activating responsive sequence, TAR. Tat seems to be a transcriptional control molecule in HTLV-1, but is both that and a post-transcriptional regulatory molecule in HTLV-III. Deletion mutants in the tat gene are incapable of prolific replication and exhibit no cytopathic effects in T4 cell lines.
 In addition to the
 9.4 kb genomic mRNA, subgenomic mRNAs of 7.4, 5.5, 5.0, 4.3, 2.0 and 1.8 kb have been detected.

FEATURES

source

Location/Qualifiers
 1. .8932

/organism="Human immunodeficiency virus 1"

/mol_type="genomic RNA"

/db_xref="taxon:11676"

prim_transcript <1..>8932

/note="genomic mRNA"

prim_transcript <1..>8932

/note="tat, rev, nef subgenomic mRNA"

Intron

CDS

/note="tat, rev, nef subgenomic mRNA intron 1"

/note="gag polypolypeptide precursor"

/codon_start=1

/protein_id="AA44201.1"

/translation="MGARASVLSGGLDEDRKIRLRGGKKYKRLKHTWASRELEF

AVNPGLTSEGCRLIGQLPISLOTSELSRLYNVALVCVHORIEIKRLEAD

KLEEDONKSKRAQAQAADTGHSSVSNPPIYONIGOVHQAISPTANVAWE

EKAESPVEIPMFALSEGAPDILMTNLNTVGHOAMQMKETINDEAAEMWVHV

HAGPIAFGQREPRSDIAGTSTLQIOMNTNPIPIVGEIYKRWIILGLNKIVM

YSPSTLIDIRGKRPDRDYVDYFRTLRAPQASOEKMMTEFLIVONNPCKTLL

KAQSPATLEEMMTACOGVGGPRHKAIVLAASQYNTFTIMQKRNPNQKRMVC

PNCGKSGHTRKNCRAPRKCGMCGKGHOMKQDTEANFGLKIMPSYGRPNLQ

SREPTAPRPLQSRPEPTAPPESSFRSGVETPTTPQKQEPIDKELPLTSLRSLFND

SSQ"

1407..4454

/gene="pol"

<1407..4454

gene

CDS

/note="ORF for pol polypeptide containing stop codons"

/codon_start=1

/product="pol polypeptide"

/protein_id="AA44198.1"

/db_xref="GI:326385"

/translation="FRREDLAFIQAQKARESSQETRANSPPTISSEQTRANSPTRREIQ

VWGRDNSSPEAGADROGVTFNFPQITTMORPLVITIKIGQLKEALLDGDADTVLE

EMSLPGRMKPKMIGIGGIFKVRQYDQILIEIGHKAIGVLVGPVNTIGRLTLQ

IGCTLNPIPIETVPPVPLKPGMDGPKVOMPLTEERIKALVEICTEMEGKRSKIG

PENDYNTPPAIKKDSTKMRKLINPRELNKRPDDEWVLOGIPIHAPGLKKKSVTL

DNGDAVSPLDDEFRKYTAFTTIPSTINNEPGRFYQYNYVPOCHKSPALFOSMRTI

LEPPKONPDIIVITQIMDDLYVGSDEIGQRTKIEELRHLRLKMGILTPDKHOKRP

PTLMGELHPRDWYTOPIVLPKDSWTVNDIQKLVGLKLMASQIYPIGVRLQCKLL

RGTALTEVPLTEAELELAENREILKEPVHGYVDSKDLAELQKQOGQYVQI

YOEPRKLTGKTAARMGATNDVOKLTEAVOKTTTESIYIMGKTPKFKIPIKQEME

TMTEVWQATWIPMEFEVNTPLVYKLMYOLEKEPIVGAETFEVYDGAANRETKGKAY

VTNKGKQVYPLNTNOKTELOAIYIALADSGLEVNIYDSQALCIIOAQDPKSS

ELVNQILEQIKKQVYLAWPAPAKGIGGNEQYKLVASIGIRILFLDQIDKQDEHE

KYSNMRRMASDNLPRVPAKEIVASCDQLGKGAHQVDSPPGIMQDLCHEK

VILVAIVASGYIEAEVIPAETQETAYFLKLAGRPVPTIHTDNGSNTFSVLEKA

CNMAAGIKOEFGIYPNPOSGVESMKNELKKIIGOVDAQEHKTAQVMAVFIHNEFR

KGIGGYSAGERTIVDITADTIOIKELQKOTKIONPEVYVYRDSRNPJMKRPALLMG

EGAVVIDNSDIDIVPRRRKAKITRDYKQAGADCAASRODED"

4399..4977

CDS

/note="vif protein"

/codon_start=1

/protein_id="AA44202.1"

/db_xref="GI:326389"

/translation="MKNRQVMTVMQVDRMRITKSLVYKHHMYVSGKANGMYRIHY

ESPHRISSEVHIPLGDARLVITTYWGLHGERDWHLGQVSIEMRRKRYSTVDPEL

ADQLHLVYDFDCSSDAIRKALLGHIVSPCEYQAGHNKVGSLQYIALAALITPKKIK

CDS

PLPSTYKLTEDRMNRPQKTKGHSHTNGH"
 4917..5153

/note="vpr protein"

/codon_start=1

/protein_id="AA44203.1"

/db_xref="GI:326390"

/translation="MEQAPEDQSPQREPHNWTLELLELKEAVRHFPRIWHLGIGQ

HIYETGDMAGVEAIRILQQLFIHFQWVST"

join(5189..5403,7734..7779)

/note="tat protein"

/codon_start=1

/protein_id="AA44199.1"

/db_xref="GI:326386"

/translation="MEPVDRPLEPMKHPGSPQPTACTNCCYCKKCFHCQVFTKALG

ISYGRKKRORRRPPOGSDTHVSLSKQPTSGRGPPTGPK"

<5189..5403

/note="tat protein, (first expressed exon)"

/number=2

join(5328..5403,7734..8008)

/note="rev protein"

/codon_start=1

/protein_id="AA44200.1"

/db_xref="GI:326387"

/translation="MAGRSGDSEDDLKAVRLIKFLYQSNPPNPECTQARRRRR

WRERQRIHSISERILSTYLGRSABEVPLQPLERLTLDNEDCGTSGTGVGSPQI

LVSPTVLESQAKE"

<5328..5403

/note="rev protein, (first expressed exon)"

/number=2

5404..7733

/note="tat cds intron 2"

5404..7733

/note="rev cds intron 2"

5404..7733

/note="tat, rev, nef subgenomic mRNA intron 2"

5420..5665

/note="vpu protein"

/codon_start=1

/protein_id="AA44204.1"

/db_xref="GI:326391"

/translation="MQPIQAIYALVVAITAIYVMSIYIIEFRKILRQKIDRLDR

LIERAEDSGNESGEISALVEKGVEMGHAPWVDL"

5580..8150

/note="envelope polypeptide"

/codon_start=1

/protein_id="AA44205.1"

/db_xref="GI:326392"

/translation="MRVKEKYQHLMRGWRMGVTLGLMLICATEKLVNTYVYQVY

WKBATYTLFCASAKAYDTEVHNVMAIHAQVPPDPNPOEVVLAIVNENFMWNRNDVE

QMHEDIISLMDQSLPCVKLTPLCVSLKCDLKNDRNTNSSGRLMEKEIKNSFN

ISHSIRGKVOKEVAFYKLDIIPIDNDTSTYTLSCNTSVITQACPKVSEPIPIHC

APAGPGLKNNKNTFNGTCTVNVYCHGIRBPVYQDLLNGSLAEVEEYIRSAN

FTPMNAITIIYQLODSVEINCTPRNNNTFRKIRIRORGBRFPYITGKIGNRRQAHIS

RAWMNNTLKQIDSKLRQEGNNKTIIFKQSSGDPETVHSFNGGDEFFYCNSTOLFN

STFNSTNWSKSGNNTEGSDTITLPKRIQIIMMOQEVKAMAPPSQIRKSSNIT

GLLTLRKQSGNNESEIFRGGGDMDMNMSLEYKKVYVIEELGVAPTKARRVOR

EKAVAGIGALFIEFLGAGSTWGAASWTIVQAROLSGIVOOONNLTRAIEQOHL

QIVWQIKQIOLAEILIAVERYLKQOOLGIMGCGKLICTPAVPMNASMSKSIQIWN

NMTWMEHREINNTSILHSLESQNOQKRNQDELELDKMASLKNWPIITWMTYI

KLPIMVIGVGLIRIVFAVLSVNRKQGSPLSFQTHLPITPGPDRPGIEEGGER

DRORSIRLVNGSLALIMDDIRSLCFESHYLRDLILVIRIVELGRGEWALRYMN

LIVYQSELKNSAVILLNATAIVAEQTDVIVVQAGVRAIRHIPPRIKQIGERILL

"

7734..>8008

/note="rev protein"

/number=3

7734..>7779

/note="tat protein"

/number=3

8152..8523

/note="nef protein, (first expressed exon; premature

termination)"

/number=3

8152..8523

/number=3


```
/codon_start=1
/protein_id="AA44206.1"

Query Match      41.7%; Score 2597.8; DB 14; Length 8932;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 2639; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

OY 1291 CCATGAGGTGAGAGAGAGATATGACGACCTGTGTGAGATGGGGGTGGAATGGGCACCA 1350
    |||||||
Db 5578 CAATGAGGTGAGAGAGAGAAATATGACGACCTGTGTGAGATGGGGGTGGAATGGGCACCA 5637

OY 1351 TGTCTCTGGGATATGTGATCTGTACTGTCTACAGAAAATTTGGGTCAACCGCTATT 1410
    |||||||
Db 5638 TGTCTCTGGGATGTGTATGTATCTGTACTGTCTACAGAAAATTTGGGTCAACCGCTATT 5697

OY 1411 ATGGGATCTGTGTGAGAGAGACACCACTCTATTTTGTGCATCAGATGCTAAAG 1470
    |||||||
Db 5698 ATGGGATCTGTGTGAGAGAGACCACTCTATTTTGTGCATCAGATGCTAAAG 5757

OY 1471 CATATGATACAGAGGTACATATATGTGTGGGCACACATGCTGTACCCACAGACCCA 1530
    |||||||
Db 5758 CATATGATACAGAGGTACATATATGTGTGGGCACACATGCTGTACCCACAGACCCA 5817

OY 1531 ACCCACAAGAGTATGTGTAATGTGACAGAAAATTTTAACTGTGAAAATGACA 1590
    |||||||
Db 5818 ACCCACAAGAGTATGTGTAATGTGACAGAAAATTTTAACTGTGAAAATGACA 5877

OY 1591 TGGTGAACACATGATGATATATCACTTTATGGGATCAAAAGCCTAAAGCCTATG 1650
    |||||||
Db 5878 TGGTGAACACATGATGATATATCACTTTATGGGATCAAAAGCCTAAAGCCTATG 5937

OY 1651 TAAATTTACCCACTCTGTGTAGTTTAAAGTGACATGATTTGAAGATGTACTATA 1710
    |||||||
Db 5938 TAAATTTACCCACTCTGTGTAGTTTAAAGTGACATGATTTGAAGATGTACTATA 5997

OY 1711 CCAATAGTAGAGGGGAGAAATGATTAATGAGAAAAGAGATAAAAAAGTCTCTTTCA 1770
    |||||||
Db 5998 CCAATAGTAGAGGGGAGAAATGATTAATGAGAAAAGAGATAAAAAAGTCTCTTTCA 6057

OY 1771 ATATCAGACAAAGCTTAAGAGATTAAGTGCAGAAAAGATATGCAATCTTTATTAACCTG 1830
    |||||||
Db 6058 ATATCAGACAAAGCTTAAGAGATTAAGTGCAGAAAAGATATGCAATCTTTATTAACCTG 6117

OY 1831 ATATAGTACCAATAGATA-----TACCAGCTATAGTTGTAAAGTGTAAACCTGAG 1884
    |||||||
Db 6118 ATATAGTACCAATAGATA-----TACCAGCTATAGTTGTAAAGTGTAAACCTGAG 6177

OY 1885 TCATTTACAGAGGCTGTCCAAAGGTATCTTTTGAGCCAAATCCCATATATTGTGCCC 1944
    |||||||
Db 6178 TCATTTACAGAGGCTGTCCAAAGGTATCTTTTGAGCCAAATCCCATATATTGTGCCC 6237

OY 1945 CGGCTGCTTTTCCGATTTTAAATGTAAATTAAGAGCTTCAATGGAACAGACCATGTA 2004
    |||||||
Db 6238 CGGCTGCTTTTCCGATTTTAAATGTAAATTAAGAGCTTCAATGGAACAGACCATGTA 6297

OY 2005 CAAATGTGACAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2064
    |||||||
Db 6298 CAAATGTGACAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6357

OY 2065 TGTTAATATGAGAGTCTGACAGAAAGATGTAGTAATTAATGATCTGCCAATTTACAGACA 2124
    |||||||
Db 6358 TGTTAATATGAGAGTCTGACAGAAAGATGTAGTAATTAATGATCTGCCAATTTACAGACA 6417

OY 2125 ATGCTAAAACCATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2184
    |||||||
Db 6418 ATGCTAAAACCATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6477

OY 2185 ACAACAATACAGAAAAAGATATCCGTATCCAGAGGGAGCCAGGGAGACATTTGTTACAA 2244
    |||||||
Db 6478 ACAACAATACAGAAAAAGATATCCGTATCCAGAGGGAGCCAGGGAGACATTTGTTACAA 6537

OY 2245 TAGGAAAAATAGAAATATGAGACAAAGCAATTTGTAACTTAGAGACAAAATGGAATG 2304
    |||||||
```

```
Db 6538 TAGGAAAAATAGAAATATGAGACAAAGCAACATTTGTAACTTAGTAGACAAAATGGAATA 6597
OY 2305 CCACCTTAAACAGATTAAGCAATTAAGAAACAAATTTGGAAATTAATAAACAAATAA 2364
    |||||||
Db 6598 ACACCTTAAACAGATTAAGCAATTAAGAAACAAATTTGGAAATTAATAAACAAATAA 6657

OY 2365 TCTTTAACCAATCCACAGAGGGAGCCAGAAATGTAAACGACAGTTTAAATTTGTGAG 2424
    |||||||
Db 6658 TCTTTAACCAATCCACAGAGGGAGCCAGAAATGTAAACGACAGTTTAAATTTGTGAG 6717

OY 2425 GGAATTTTCTACTGTATTAATTCACACAACTGTTTAATAGTACTGTTTAAATAGTACTT 2484
    |||||||
Db 6718 GGAATTTTCTACTGTATTAATTCACACAACTGTTTAATAGTACTGTTTAAATAGTACTT 6777

OY 2485 GGAGTACTGAAGGCTCAATAATTAACCTGAAGAGTGAACAAATCACCCTCCATGACAGA 2544
    |||||||
Db 6778 GGAGTACTGAAGGCTCAATAATTAACCTGAAGAGTGAACAAATCACCCTCCATGACAGA 6837

OY 2545 TAAACCAATTTAATAACATGTGGGAGAAAGTGAAGAAAAGCAATGATGCCCCCATCA 2604
    |||||||
Db 6838 TAAACCAATTTAATAACATGTGGGAGAAAGTGAAGAAAAGCAATGATGCCCCCATCA 6897

OY 2605 GTGACAAATTAAGATGTTTCATCAATTAATTAAGTGGCTCTATTAAACAAGAGATGGTGTA 2664
    |||||||
Db 6898 GTGACAAATTAAGATGTTTCATCAATTAATTAAGTGGCTCTATTAAACAAGAGATGGTGTA 6957

OY 2665 ATAACAATAATGGGTCCAGATCTTCAACACTGGAGAGGCGATATGAGGCAATTTGGA 2724
    |||||||
Db 6958 ATAACAATAATGGGTCCAGATCTTCAACACTGGAGAGGCGATATGAGGCAATTTGGA 7017

OY 2725 GAAGTAATTTATTAATTAATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 2784
    |||||||
Db 7018 GAAGTAATTTATTAATTAATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 7077

OY 2785 AGGCAAGAAGAGAGTGTGACAGAGAAAAAAGAGAGAGTGGGAATAGAGCTTTGTTCC 2844
    |||||||
Db 7078 AGGCAAGAAGAGAGTGTGACAGAGAAAAAAGAGAGAGTGGGAATAGAGCTTTGTTCC 7137

OY 2845 TTGGTCTTTGGGACACAGAGAACCACTATAGGCTGACACCTCAATGACCTGAGAGGTAC 2904
    |||||||
Db 7138 TTGGTCTTTGGGACACAGAGAACCACTATAGGCTGACACCTCAATGACCTGAGAGGTAC 7197

OY 2905 AGGCGAGACAAATTAATGCTATATATGATGATGATGATGATGATGATGATGATGATGATG 2964
    |||||||
Db 7198 AGGCGAGACAAATTAATGCTATATATGATGATGATGATGATGATGATGATGATGATGATG 7257

OY 2965 AGGCGCAACAGACATCTGTTGCAACTGACAGTCTGGGGCATCAACAGCTCCAGGCAAGAA 3024
    |||||||
Db 7258 AGGCGCAACAGACATCTGTTGCAACTGACAGTCTGGGGCATCAACAGCTCCAGGCAAGAA 7317

OY 3025 TCCTGGCTGTGGAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3084
    |||||||
Db 7318 TCCTGGCTGTGGAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 7377

OY 3085 GAAAACCTATTTGACACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3144
    |||||||
Db 7378 GAAAACCTATTTGACACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7437

OY 3145 AACAGATTTGGAATTAACATGACCTGATGATGATGATGATGATGATGATGATGATGATGATG 3204
    |||||||
Db 7438 AACAGATTTGGAATTAACATGACCTGATGATGATGATGATGATGATGATGATGATGATGATG 7497

OY 3205 GCTTAATACCTCTTAATTAAGAAATGCAAAACACAGAAAGAAATGAACAGAAAT 3264
    |||||||
Db 7498 GCTTAATACCTCTTAATTAAGAAATGCAAAACACAGAAAGAAATGAACAGAAAT 7557

OY 3265 TATTGGAATTTAGATTAATGAGCAAGTTTGTGGAATTTGTTTAACTAAACAAATTTGGCTGT 3324
    |||||||
Db 7558 TATTGGAATTTAGATTAATGAGCAAGTTTGTGGAATTTGTTTAACTAAACAAATTTGGCTGT 7617

OY 3325 GGTATATATAAATTTATCAATAATGATAGTAGAGGCTGTGTAGGTTTAAAGAAATAGTTTGTG 3384
    |||||||
Db 7618 GGTATATATAAATTTATCAATAATGATAGTAGAGGCTGTGTAGGTTTAAAGAAATAGTTTGTG 7677
```

QY 3385 CTGTACTTCTATAGTAATAGAGTAGGCGAGATATTCACATTATCGTTTCAGACC 3444
|||||
Db 7678 CTGTACTTCTGTAGTAGTAATAGAGTAGGCGAGATATTCACATTATCGTTTCAGACC 7737
|||||
QY 3445 ACCTCCCAATCCGAGGGGACCCGACAGGCCGAGAGATAGAGAAGAGGTGGAGAGA 3504
|||||
Db 7738 ACCTCCCAATCCGAGGGGACCCGACAGGCCGAGAGATAGAGAAGAGGTGGAGAGA 7797
|||||
QY 3505 GAGACAGAGACAGATCCATTGATTAAGAACGGATCCCTTAGCACTTATCTGGAGCATC 3564
|||||
Db 7798 GAGACAGAGACAGATCCATTGATTAAGAACGGATCCCTTAGCACTTATCTGGAGCATC 7857
|||||
QY 3565 TGCAGACCTGTGCTCTTACGTACACCGCTTAGAGACTTACTCTTGATTGTAAAGA 3624
|||||
Db 7858 TGCAGACCTGTGCTCTTACGTACACCGCTTAGAGACTTACTCTTGATTGTAAAGA 7917
|||||
QY 3625 GGATTGTGAACTTCTGGAGCGCAGGGGGTGGAAAGCCCTCAATATTGGTGAATCTCC 3684
|||||
Db 7918 GGATTGTGAACTTCTGGAGCGCAGGGGGTGGAAAGCCCTCAATATTGGTGAATCTCC 7977
|||||
QY 3685 TACAGTATGTGAGTCAGAGAACTAAGAAATAGTGTCTTAATCTCAATGCCACAGCA 3744
|||||
Db 7978 TACAGTATGTGAGTCAGAGAACTAAGAAATAGTGTCTTAATCTCAATGCCACAGCA 8037
|||||
QY 3745 TAGCAGTAGCTGAGGGGACAGATAGGTTATAGAAGTATTACAGCAGCTTATAGAGCTA 3804
|||||
Db 8038 TAGCAGTAGCTGAGGGGACAGATAGGTTATAGAAGTATTACAGCAGCTTATAGAGCTA 8097
|||||
QY 3805 TTGCGCACATACCTAGAGAATAAGACAGGGCTTGAAAGGATTTGCTATATAGATGGGT 3864
|||||
Db 8098 TTGCGCACATACCTAGAGAATAAGACAGGGCTTGAAAGGATTTGCTATATAGATGGGT 8157
|||||
QY 3865 GGCAGAGTGTCAAAAGTATAGTGTATGGATGGCTCTGAAGGAAAGAAATGAGAGA 3924
|||||
Db 8158 GGCAGAGTGTCAAAAGTATAGTGTATGGATGGCTCTGAAGGAAAGAAATGAGAGA 8217
|||||
QY 3925 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGTATCTCGAGATCTAGA 3971
|||||
Db 8218 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGCAGTATCTCGAGACCTAGA 8264
|||||

Search completed: September 17, 2003, 14:10:44
Job time : 21922 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2003, 06:21:59 : Search time 1469 Seconds
(without alignments)
11446.426 Million cell updates/sec

Title: US-09-913-159A-10

Perfect score: 6229

Sequence: 1 ctgcagcgcctctgacgcgc.....attccccgaagaatgcac 6229

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N.Geneseq_19Jun03:*

- 1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
- 25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	6229	100.0	6229	AAA97934	HIV-1 env DNA frag
2	2676.2	43.0	9709	AAO96140	HIV-1 NT4-3 genome
3	2676.2	43.0	9709	AAV81871	Nucleic acid seque
4	2676.2	43.0	9709	AAA97926	HIV-1 env DNA. Hu
5	2676.2	43.0	9709	AAA40298	HIV-1 viral protei
6	2676.2	43.0	9709	AAZ88127	HIV-1 NV5/IIIB rec
7	2676.2	43.0	9709	ABO76897	HIV-1 NT4-3 gp120
8	2676.2	43.0	9709	ABSS7823	HIV-1 genome clone

9	2669.8	42.9	9709	13	AAO22488	HIV-1 proviral clo
10	2658.6	42.7	9965	24	ABO81389	HIV-1 R8 genomic D
11	2636.6	42.3	15581	18	AAV14355	Plasmid pNLNG11 e
12	2607.4	41.9	9213	20	AAV74271	HTLV-III genomic D
13	2599.4	41.7	9213	20	AAV81866	Nucleic acid seque
14	2598.2	41.7	3156	7	AAV60128	Sequence of the en
15	2597.8	41.7	8932	22	AAH76385	Nucleotide sequenc
16	2597.8	41.7	8933	22	AAF60542	Plasmid BH10 RNA t
17	2597.8	41.7	8933	22	AAF60543	IAC-Ascr pseudo ta
18	2597.8	41.7	8933	22	AAF60548	IAC-Bscr pseudo ta
19	2597.8	41.7	9425	14	AAO45912	HTLV-III DNA (Clon
20	2597.8	41.7	9425	14	AAO45919	HTLV-III DNA (Clon
21	2597.8	41.7	9749	14	AAO45922	HTLV-III DNA (Clon
22	2596.2	41.7	9213	7	AAH60288	Sequence of the HT
23	2593	41.6	9181	22	AAF24377	Human immunodefici
24	2593	41.6	9181	25	AAD49652	Human immunodefici
25	2593	41.6	9181	25	AAO53720	HTLV-III DNA sequenc
26	2591.4	41.6	9427	14	AAO45920	HTLV-III DNA (Clon
27	2591.4	41.6	9745	7	AAH60240	HTLV-III virus (HI
28	2590.8	41.6	2945	16	AAH05127	HIV virus-1 gp160
29	2572.2	41.3	9609	24	AAH25515	Human immunodefici
30	2569.6	41.3	9193	24	AAH49920	Lymphadenopathy-as
31	2561.6	41.1	9193	7	AAH60365	Sequence of LAV vi
32	2560.4	41.1	2562	22	AAH20891	HIV gp41 DNA seq I
33	2558.4	41.1	12494	20	AAH83206	Vector pHP-1 compr
34	2558.4	41.1	12494	22	AAF85623	lentiviral vector
35	2557	41.0	2565	22	AAF56366	HIV gene. Human 1
36	2545.4	40.8	9088	7	AAH60476	Sequence of Lympha
37	2476	39.7	10006	24	AAD25514	Simian-Human immun
38	2454	39.4	2607	7	AAH60077	DNA sequence of th
39	2444.4	39.2	4020	8	AAH71016	Sequence of LAV/HT
40	2333.2	37.5	2641	13	AAO30970	Non-cleavable, sol
41	2310.4	37.1	2553	9	AAH80949	HIV protein HT6.
42	2234.4	35.9	8298	19	AAV18096	pPCIDHAP tricitstro
43	2206.6	35.4	3806	20	AAH04767	Env gene of the BA
44	2205.6	35.4	3807	18	AAH58551	Human immunodefici
45	2197.4	35.3	8560	21	AAA93981	HIV vector pHAD0ZE

ALIGNMENTS

RESULT 1	AAA97934	standard; DNA: 6229 BP.
AAA97934		
AC	AAA97934:	
XX		
XX	26-JAN-2001	(first entry)
DT		
XX		
DE	HIV-1 env DNA fragment #2 in plasmid pBSCenvATC.	
XX		
KW	Viral protein; vaccine; anti-viral; anti-HIV; therapy; infection;	
KW	env protein; envelope protein; ds.	
XX		
OS	Human immunodeficiency virus type 1.	
XX		
PN	WO200047223-A2.	
XX		
PD	17-AUG-2000.	
XX		
PE	03-DEC-1999;	99WO-EP09759.
XX		
PR	12-FEB-1999;	99DE-1007485.
XX		
PA	(STRA-) STRATHMANN & CO AG.	
XX		
PI	Schreiber M;	
XX		
DR	WPI; 2000-549084/50.	
XX		
PT	Viral vaccine comprises a mixture of protein sequence variants of a	
PT	single viral protein, which is useful for prevention and therapy of	

PT viral infections, especially HIV, in humans
XX
PS Claim 28; Page 76-78; 79pp; German.
XX
CC This invention describes a novel protein vaccine which comprises a
CC mixture of viral proteins, characterized in that the molecules are
CC sequence variants of a single viral (partial) protein. The products of
CC the invention have anti-viral and anti-HIV activity. Mixtures of
CC structurally different viral proteins, that are sequence variants of a
CC single protein are useful for production of vaccines for the prevention
CC and/or therapy of viral infections in humans. The vaccines are especially
CC useful for prevention and/or therapy of human immunodeficiency virus
CC (HIV) infection in humans. The vectors and host cells are useful for the
CC expression of the protein/DNA mixtures, which are also useful for the
CC preventing and/or therapy of viral infection. This sequence encodes
CC a fragment of the human immunodeficiency virus (HIV-1) envelope (env)
CC protein which is used to illustrate the method of the invention.
XX
S0 Sequence 6229 BP; 1795 A; 1332 C; 1518 G; 1584 T; 0 other;

Query Match 100.0%; Score 6229; DB 21; Length 6229;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACGGCCCTGTAGCGGCGCATTAAGCGCGGGGTGTGTGTTACGGCAGCGTGA 60
DB 1 CTGACGGCCCTGTAGCGGCGCATTAAGCGCGGGGTGTGTGTTACGGCAGCGTGA 60
QY 61 CCGGTACACTTGGCCAGGCGCCCTACGCCGCTCTTCCTTCCTTCCTTCCTTCG 120
DB 61 CCGGTACACTTGGCCAGGCGCCCTACGCCGCTCTTCCTTCCTTCCTTCCTTCG 120
QY 121 CCAGCTTGCAGCGGCTTCCCGCTCAAGCTTAATCGGGGCTCCCTTAAAGGTTCCGAT 180
DB 121 CCAGCTTGCAGCGGCTTCCCGCTCAAGCTTAATCGGGGCTCCCTTAAAGGTTCCGAT 180
QY 181 TTATGCTTTAAGCGCACCTCGACCCCAAAAATTGATAGGATGATGTTACAGTAGTG 240
DB 181 TTATGCTTTAAGCGCACCTCGACCCCAAAAATTGATAGGATGATGTTACAGTAGTG 240
QY 241 GGCCATGGCCCTGATPAGAGGTTTGGCCCTTGGAGTGGAGTCCAGCTTCTTATA 300
DB 241 GGCCATGGCCCTGATPAGAGGTTTGGCCCTTGGAGTGGAGTCCAGCTTCTTATA 300
QY 301 GTGACCTCTTGTTCACAACTGGAACAACACTCAACCTATCTCGTCTATCTTTGATT 360
DB 301 GTGACCTCTTGTTCACAACTGGAACAACACTCAACCTATCTCGTCTATCTTTGATT 360
QY 361 TATPAGGATTTTCCGATTTGGCTTATGTTAAAAATGAGCTGATTTAACAAAAAT 420
DB 361 TATPAGGATTTTCCGATTTGGCTTATGTTAAAAATGAGCTGATTTAACAAAAAT 420
QY 421 TTACGGGAATTTTAAACAAATATTAAAGCTTACAAATTCATTCGACATTCAGGCTCG 480
DB 421 TTACGGGAATTTTAAACAAATATTAAAGCTTACAAATTCATTCGACATTCAGGCTCG 480
QY 481 CAACCTTTGGGAAGGCGATCGTGGCGCTTCTCCATATTACGCACTGGGGAAG 540
DB 481 CAACCTTTGGGAAGGCGATCGTGGCGCTTCTCCATATTACGCACTGGGGAAG 540
QY 541 GGGATGTGTGCAAGGGGATTAAGTTGGTAAACCCAGGGTTTCCAGTACAGACGTTG 600
DB 541 GGGATGTGTGCAAGGGGATTAAGTTGGTAAACCCAGGGTTTCCAGTACAGACGTTG 600
QY 601 TAAAGAGCGGCGGTAGAGGTCTAGTTATTAATGAATCAATTAAGGGGTCTATAGTT 660
DB 601 TAAAGAGCGGCGGTAGAGGTCTAGTTATTAATGAATCAATTAAGGGGTCTATAGTT 660
QY 661 CATAGCCCATATATAGAGTTCCGGTGTACATTAACCTTACGTAATATGCGCCGCTG 720
DB 661 CATAGCCCATATATAGAGTTCCGGTGTACATTAACCTTACGTAATATGCGCCGCTG 720
QY 721 CCGCCCAAGACGCCCGCCCATTTAGCTCAATTAATGAAGTATGTTCCCATAGTAACGCCA 780

DB 721 CCGCCCAAGACGCCCGCCCATTTAGCTCAATTAATGAAGTATGTTCCCATAGTAACGCCA 780
QY 781 ATAGGACTTTCATTTAGCTCAATGSGTGAAGTATTTAGCGTAACTGCCACTTGGCA 840
DB 781 ATAGGACTTTCATTTAGCTCAATGSGTGAAGTATTTAGCGTAACTGCCACTTGGCA 840
QY 841 GTACATCAAGGTATTCATATGCAAGTACGCGCCCTTATGACGTTAACTGAGGTAATGG 900
DB 841 GTACATCAAGGTATTCATATGCAAGTACGCGCCCTTATGACGTTAACTGAGGTAATGG 900
QY 901 CCCGCTGCAATTAATGCGCATACATGACCTTATGAGGACTTCTTACCTTGGCAGTATAC 960
DB 901 CCCGCTGCAATTAATGCGCATACATGACCTTATGAGGACTTCTTACCTTGGCAGTATAC 960
QY 961 TACGTATTAATGATCATGCTTATTTACATGATGATGATGATGATGATGATGATGATG 1020
DB 961 TACGTATTAATGATCATGCTTATTTACATGATGATGATGATGATGATGATGATGATG 1020
QY 1021 GGATAGGCGTTTGAATCAGGCGGATTTCCAAAGTCTCCAGCCCATTTAGCGTCAATGGAGT 1080
DB 1021 GGATAGGCGTTTGAATCAGGCGGATTTCCAAAGTCTCCAGCCCATTTAGCGTCAATGGAGT 1080
QY 1081 TTGTTTGGCACCAAAATCAACGGGACTTTCACAAATGTGTAACAACTCCGCCCATTTG 1140
DB 1081 TTGTTTGGCACCAAAATCAACGGGACTTTCACAAATGTGTAACAACTCCGCCCATTTG 1140
QY 1141 ACGCAATGAGCGGTAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1200
DB 1141 ACGCAATGAGCGGTAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1201 AACCGTACATGCGCTGAGAGCGCATCCAGCTGTTTACCTCCATAGAACACACCGG 1260
DB 1201 AACCGTACATGCGCTGAGAGCGCATCCAGCTGTTTACCTCCATAGAACACACCGG 1260
QY 1261 GACAAATTCAGCTCGTACCGTGCAGCGCCACATGAGAGTGAAGAAAGTATCAGCACT 1320
DB 1261 GACAAATTCAGCTCGTACCGTGCAGCGCCACATGAGAGTGAAGAAAGTATCAGCACT 1320
QY 1321 TGTGGAATGAGGAGTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
DB 1321 TGTGGAATGAGGAGTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
QY 1381 CTACAGAAAAATGTGGGTACCGCTCTATTATGGGATCCGTGTTGGAAGGAAGCAACCA 1440
DB 1381 CTACAGAAAAATGTGGGTACCGCTCTATTATGGGATCCGTGTTGGAAGGAAGCAACCA 1440
QY 1441 CCACCTATTTTGTGATCAGATGCTAAAGCATATGATGATGATGATGATGATGATGATG 1500
DB 1441 CCACCTATTTTGTGATCAGATGCTAAAGCATATGATGATGATGATGATGATGATGATG 1500
QY 1501 CCACACATGCGCTGTGATCCACAGAGCCCAACCCACAGAGAGTATGTTGTAATGTGA 1560
DB 1501 CCACACATGCGCTGTGATCCACAGAGCCCAACCCACAGAGAGTATGTTGTAATGTGA 1560
QY 1561 CAGAAAAATTTTAAACATGTGGAAGAAATGACATGTGTAAGACAGATGATGATGATGAT 1620
DB 1561 CAGAAAAATTTTAAACATGTGGAAGAAATGACATGTGTAAGACAGATGATGATGATGAT 1620
QY 1621 GTTTATGAGATCAAAAGCTTAAGCATGTGTAAATTAACCCACATCTGTGTTAGTTAA 1680
DB 1621 GTTTATGAGATCAAAAGCTTAAGCATGTGTAAATTAACCCACATCTGTGTTAGTTAA 1680
QY 1681 AGTGCACTGATTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
DB 1681 AGTGCACTGATTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
QY 1741 AGAAGAGAGATTAATAAACTGCTTTTCAATATCAGACACACATTAAGATTAAGGTGC 1800
DB 1741 AGAAGAGAGATTAATAAACTGCTTTTCAATATCAGACACACATTAAGATTAAGGTGC 1800
QY 1801 AGAAGATATGCAATCTTTTATTAATCTGATATGATGATGATGATGATGATGATGATGAT 1860
DB 1801 AGAAGATATGCAATCTTTTATTAATCTGATATGATGATGATGATGATGATGATGATGAT 1860

Dd 1801 AGAAGAATATGATCTCTTTTATATAAAGTATAGTACCAATAGATATATACAGACTATA 1860
Oy 1861 GGTTGATAGATTGTAACACCTCATGATTTACACAGGCGCTGCCAAAGGTATCTTTGAGC 1920
Dd 1861 GGTGGATAAGATTGTAACACCTCATGATTTACACAGGCGCTGCCAAAGGTATCTTTGAGC 1920
Oy 1921 CAATTTCCCATCATTAATTTGTCGCCGCGCTGTTTGGGATCTCTAAATGATTAATATAAGA 1980
Dd 1921 CAATTTCCCATCATTAATTTGTCGCCGCGCTGTTTGGGATCTCTAAATGATTAATATAAGA 1980
Oy 1981 CTTTCATGAGAACAGGACCATGTACAAATGTACAGACAGTACAAATGTACACATGGAATCA 2040
Dd 1981 CTTTCATGAGAACAGGACCATGTACAAATGTACAGACAGTACAAATGTACACATGGAATCA 2040
Oy 2041 GGCAGAGTATGTCACAACTGCTGTTAAATGGCAGCTAGCAGAGAGAAATGTAGTAA 2100
Dd 2041 GGCAGAGTATGTCACAACTGCTGTTAAATGGCAGCTAGCAGAGAGAAATGTAGTAA 2100
Oy 2101 TTAGATCTGCCAATTTTCAGACAAATGCTAAACCATATATAGTACAGCTGGAACATCTG 2160
Dd 2101 TTAGATCTGCCAATTTTCAGACAAATGCTAAACCATATATAGTACAGCTGGAACATCTG 2160
Oy 2161 TAGAAATTAATTTGTAACAAGACCCAAACAATACAGAGAAAAAGTATCCGTATCCAGAGGG 2220
Dd 2161 TAGAAATTAATTTGTAACAAGACCCAAACAATACAGAGAAAAAGTATCCGTATCCAGAGGG 2220
Oy 2221 GACACAGGAGAGCAATTTGTTTCAATAGAGAAAAATAGAAATAGACAAACACATCTGTA 2280
Dd 2221 GACACAGGAGAGCAATTTGTTTCAATAGAGAAAAATAGAAATAGACAAACACATCTGTA 2280
Oy 2281 ACATTTAGTAGAGCAAAATGGAATGCCACTTTAAAAAGATAGCTAGCAAAATTAAGAGAAC 2340
Dd 2281 ACATTTAGTAGAGCAAAATGGAATGCCACTTTAAAAAGATAGCTAGCAAAATTAAGAGAAC 2340
Oy 2341 AATTTGGAATTAATTAACAAATTAATCTTTAAACAATCTCAGAGGGGAGCCAGAAATG 2400
Dd 2341 AATTTGGAATTAATTAACAAATTAATCTTTAAACAATCTCAGAGGGGAGCCAGAAATG 2400
Oy 2401 TAAAGCACAGTTTAATTTGTTGAGGGGAAATTTTCTACTGTAATTCACACACACTGTTA 2460
Dd 2401 TAAAGCACAGTTTAATTTGTTGAGGGGAAATTTTCTACTGTAATTCACACACACTGTTA 2460
Oy 2461 ATATGACTTGGTTTAATTAATGTAATGTAAGTACTGAAGGTCAAATTAACCTGAGAGAAATG 2520
Dd 2461 ATATGACTTGGTTTAATTAATGTAATGTAAGTACTGAAGGTCAAATTAACCTGAGAGAAATG 2520
Oy 2521 ACACAATCACACTCCCATGAGAAATTAACAAATTTATTAACATGTGGCAGAGAAATGAGAA 2580
Dd 2521 ACACAATCACACTCCCATGAGAAATTAACAAATTTATTAACATGTGGCAGAGAAATGAGAA 2580
Oy 2581 AAGCAATGTATGCCCCCTCCCATGAGTGCAGTGCACAAATTTAGATTTTACTGGGC 2640
Dd 2581 AAGCAATGTATGCCCCCTCCCATGAGTGCAGTGCACAAATTTAGATTTTACTGGGC 2640
Oy 2641 TGCATTATTAACAAGAGATGTTGTTGATTAACAACAAATGGGTCCGAGATCTTCAGACCTGAG 2700
Dd 2641 TGCATTATTAACAAGAGATGTTGTTGATTAACAACAAATGGGTCCGAGATCTTCAGACCTGAG 2700
Oy 2701 GAGCGATATGAGGAGCAATTTGAGAGAGTGAATTAATTAATTAAGTACTAAATAATG 2760
Dd 2701 GAGCGATATGAGGAGCAATTTGAGAGAGTGAATTAATTAATTAAGTACTAAATAATG 2760
Oy 2761 AACCATTAGAGAGTGCACCCACCAAGGCAAGAGAGAGTGTGCAGAGAGAGAGAGAG 2820
Dd 2761 AACCATTAGAGAGTGCACCCACCAAGGCAAGAGAGAGTGTGCAGAGAGAGAGAGAG 2820
Oy 2821 CAGTGGGAATGAGAGCTTTGTTCTGTTGGGCTTGGGAGCAGAGAGAGCTATGGGCT 2880
Dd 2821 CAGTGGGAATGAGAGCTTTGTTCTGTTGGGCTTGGGAGCAGAGAGAGCTATGGGCT 2880
Oy 2881 GCACGTCAATGACGCTGACGGTACAGGCGACAGCAATTAATTTGATATAGTACGAGC 2940
Dd 2881 GCACGTCAATGACGCTGACGGTACAGGCGACAGCAATTAATTTGATATAGTACGAGC 2940

Oy 2941 AGAACAAATTTGCTGAGGGCTATTTGAGGGCGCAACAGCATCTGTGCACACTCAGAGTGGG 3000
Dd 2941 AGAACAAATTTGCTGAGGGCTATTTGAGGGCGCAACAGCATCTGTGCACACTCAGAGTGGG 3000
Oy 3001 GCATCAAAACAGCTCCAGGCAAGAAATCCGCGCTGTGGAAGATACCTTAAGAGATCAACAGC 3060
Dd 3001 GCATCAAAACAGCTCCAGGCAAGAAATCCGCGCTGTGGAAGATACCTTAAGAGATCAACAGC 3060
Oy 3061 TCCTGGGGATTTGGGGTTCCTGTGGAAGAACTATTTGCACACATGCTGTGCTTGGAAATG 3120
Dd 3061 TCCTGGGGATTTGGGGTTCCTGTGGAAGAACTATTTGCACACATGCTGTGCTTGGAAATG 3120
Oy 3121 CTAGTTGAGTAATTAATCTCTGGAACGATTTGGAAATACATGACCTGATGAGAGTGG 3180
Dd 3121 CTAGTTGAGTAATTAATCTCTGGAACGATTTGGAAATACATGACCTGATGAGAGTGG 3180
Oy 3181 ACAGAGAAATTAACAAATTTACACAACTTAATACACTCTTAAATTAAGAGAAATCCGAAACC 3240
Dd 3181 ACAGAGAAATTAACAAATTTACACAACTTAATACACTCTTAAATTAAGAGAAATCCGAAACC 3240
Oy 3241 AGCAAGAAAGAAATGACAGAAATTAATGAAATTAATTAATTTGGGCAAGTTTGTGGAATT 3300
Dd 3241 AGCAAGAAAGAAATGACAGAAATTAATGAAATTAATTAATTTGGGCAAGTTTGTGGAATT 3300
Oy 3301 GGTAAACATTAACAAATTTGGCTGTGTATTAATTAATTTTCAATATGATAGTAGAGGCT 3360
Dd 3301 GGTAAACATTAACAAATTTGGCTGTGTATTAATTAATTTTCAATATGATAGTAGAGGCT 3360
Oy 3361 TGCTAGAGTTTAAAGATTAATTTTGTGCTGCTTCTATATAGTAATAGAGTATGGCAGGAT 3420
Dd 3361 TGCTAGAGTTTAAAGATTAATTTTGTGCTGCTTCTATATAGTAATAGAGTATGGCAGGAT 3420
Oy 3421 ATTCACATTAATCGTTTCAGACCCACTCCCAATCCGAGGGGAGCCGACAGGCCGAG 3480
Dd 3421 ATTCACATTAATCGTTTCAGACCCACTCCCAATCCGAGGGGAGCCGACAGGCCGAG 3480
Oy 3481 GAATAGAAAGAAAGGTGAGAGAGAGACAGACAGATTCATTCGATTAAGTGAAGGAT 3540
Dd 3481 GAATAGAAAGAAAGGTGAGAGAGAGACAGACAGATTCATTCGATTAAGTGAAGGAT 3540
Oy 3541 CCTTAGCATTATCTGGGAGCATCTGCGAGCCCTGTGCTTCAAGTACCAACGCTTGA 3600
Dd 3541 CCTTAGCATTATCTGGGAGCATCTGCGAGCCCTGTGCTTCAAGTACCAACGCTTGA 3600
Oy 3601 GAGACTTACTCTTGAATTTGAAGAGATTTGTGAACCTTCTGGACGCGAGGGGTGGAG 3660
Dd 3601 GAGACTTACTCTTGAATTTGAAGAGATTTGTGAACCTTCTGGACGCGAGGGGTGGAG 3660
Oy 3661 CCCCTCAATATTTGGTGGAAATCTCTACAGTATTTGAGATCAGAGAACTAAAGAAATAGTCTG 3720
Dd 3661 CCCCTCAATATTTGGTGGAAATCTCTACAGTATTTGAGATCAGAGAACTAAAGAAATAGTCTG 3720
Oy 3721 TTAACCTGCTCAATGCCACAGCATAGCAGTATGAGAGGAGACAGATAGGTTATAGAG 3780
Dd 3721 TTAACCTGCTCAATGCCACAGCATAGCAGTATGAGAGGAGACAGATAGGTTATAGAG 3780
Oy 3781 TATTACAAGAGCTTATAGAGCTATTTGCCACATTAACCTTAAGAGAGAGAGGCTTTGG 3840
Dd 3781 TATTACAAGAGCTTATAGAGCTATTTGCCACATTAACCTTAAGAGAGAGAGGCTTTGG 3840
Oy 3841 AAAGGATTTTCTATTAAGATGGTGGCAAGTGTGTCAAAAGTATGATGATTTGGGCT 3900
Dd 3841 AAAGGATTTTCTATTAAGATGGTGGCAAGTGTGTCAAAAGTATGATGATTTGGGCT 3900
Oy 3901 GCTGTAAGGAGAAAGATGAGAGAGCTGAGGAGAGAGTGGGTGGAGAGAGTATCT 3960
Dd 3901 GCTGTAAGGAGAAAGATGAGAGAGCTGAGGAGAGAGTGGGTGGAGAGAGTATCT 3960
Oy 3961 CGAGATCTAGACTAGAACTAGCTTCGATTCAGACATGATAAGATACATTTGATGAGATTGG 4020
Dd 3961 CGAGATCTAGACTAGAACTAGCTTCGATTCAGACATGATAAGATACATTTGATGAGATTGG 4020

QY	4021	ACAAACCCAACTAGATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGCTAT	4080
Db	4021	ACAAACCCAACTAGAAAGCAGTGAIAAAAAATGCTTTATTTGTGAAATTTGTGCTAT	4080
QY	4081	TGCTTTATTTTAAACCATTAATACCTGCAATTAACAAAGTTTAAACAACAATTCATTTCA	4140
Db	4081	TGCTTTATTTTAAACCATTAATACCTGCAATTAACAAAGTTTAAACAACAATTCATTTCA	4140
QY	4141	TTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGAGGTTTAAAAAGCAAGTAAACCTCTA	4200
Db	4141	TTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGAGGTTTAAAAAGCAAGTAAACCTCTA	4200
QY	4201	CAATGTGTATGCTGATTTATATCTGCTCGCGGTTTCGGTGATTAACGGTGAATAAC	4260
Db	4201	CAATGTGTATGCTGATTTATGATTTATCTGCTCGCGGTTTCGGTGATTAACGGTGAATAAC	4260
QY	4261	CTGTAGACATGACAGCTCCCGGAAACGGTCAACCTTGTCGTAAAGGGAATGCCGGAGC	4320
Db	4261	CTGTAGACATGACAGCTCCCGGAAACGGTCAACCTTGTCGTAAAGGGAATGCCGGAGC	4320
QY	4321	AGACAAGCCCTCAGGGGCGCTCAGCGGGTTTGGCGGGGTGTCGGGCGCAGCCATTAAC	4380
Db	4321	AGACAAGCCCGTCAAGGGGCGCTCAGCGGGTTTGGCGGGGTGTCGGGCGCAGCCATTAAC	4380
QY	4381	CAGTCACGTACGATAGAGGAGGTATATCTGCTTAATATGCGGCATCAGAGCATTTG	4440
Db	4381	CAGTCACGTACCGATAGGAGGTATATCTGCTTAATATGCGGCATCAGAGCATTTG	4440
QY	4441	TACTAGAGTACACCATATGTGCGGGCGCGTGTGGGCTTTTCCATATAGCTCGGCC	4500
Db	4441	TACTAGAGTGCACCATATGTGCGGGCGCGTGTGGGCTTTTCCATATAGCTCGGCC	4500
QY	4501	CCTGACGAGCATCACAAAAATGCAAGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTA	4560
Db	4501	CCTGACGAGCATCACAAAAATGCAAGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTA	4560
QY	4561	TAAAGATACAGGGGTTTTCCCGGGAAGTCCCTGTGGGCTCTCTGTTCCGACCCTG	4620
Db	4561	TAAAGATACAGGGGTTTTCCCGGGAAGTCCCTGTGGGCTCTCTGTTCCGACCCTG	4620
QY	4621	CCGCTTACCGGATACCTGTCGCCCTTTCCTCTCGGGAAGCGGGGCTTTCATATAGC	4680
Db	4621	CCGCTTACCGGATACCTGTCGCCCTTTCCTCTCGGGAAGCGGGGCTTTCATATAGC	4680
QY	4681	TCACGCTGTAGGTATCTCAGTTGGGTAGGTGCTTCGCAAGCTGGGCTGTGTGCAC	4740
Db	4681	TCACGCTGTAGGTATCTCAGTTGGGTAGGTGCTTCGCTCAAGCTGGGCTGTGTGCAC	4740
QY	4741	GAACCCCCCGTTCAAGCCGACCGCTGGGCTTATCCGGTAATATGCTCTTGAATCCAAC	4800
Db	4741	GAACCCCCCGTTCAAGCCGACCGCTGGGCTTATCCGGTAATATGCTCTTGAATCCAAC	4800
QY	4801	CCGGTAAGACACGACTTATCCGCACTGGACACACCACTGGTAACAGGATTTAGCAGAGCG	4860
Db	4801	CCGGTAAGACACGACTTATCCGCACTGGACACACCACTGGTAACAGGATTTAGCAGAGCG	4860
QY	4861	AGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGGCCTTAACATAAGCTACACTAGA	4920
Db	4861	AGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGGCCTTAACATAAGCTACACTAGA	4920
QY	4921	AGGACAGTATTTGGTATCTGCTGCTGTGTAAGCCAGTTACCTTGGAAAAAGATTGGT	4980
Db	4921	AGGACAGTATTTGGTATCTGCTGCTGTGTAAGCCAGTTACCTTGGAAAAAGATTGGT	4980
QY	4981	AGCTTTGATCCGGCAAAACAACACCGCTGTGTAGCGGTGTTTTTTTGTGCAACAG	5040
Db	4981	AGCTTTGATCCGGCAAAACAACACCGCTGTGTAGCGGTGTTTTTTTGTGCAACAG	5040
QY	5041	CAGATTAGCGCAGAAAAAAGATATCAAAATCCTTTGATCTTTTTCACGGGCTC	5100
Db	5041	CAGATTAGCGCAGAAAAAAGATATCAAAATCCTTTGATCTTTTTCACGGGCTC	5100
QY	5101	GACGCTGACGTGAACGAAACTCAAGTTAAAGGATTTTGGTCATGAGATTATCAAAAAG	5160

Db	5101	GACGCTCAGTGGACGAAACCTCAGTTAAAGGATTTGGTCATGAGATTATCAAAAGG	5160
QY	5161	ATCTCCACCTAGATCCCTTTTAAATTTAAAAAAGAAGTTTAAATCAATCTAAAGATATPAT	5220
Db	5161	ATCTTCACCTAGATCCCTTTTAAATTTAAAAAAGAAGTTTAAATCAATCTAAAGATATPAT	5220
QY	5221	GAGTAAACCTTGGCTGACAGTTTACCAATGCTTAAATCAAGTAGGACCTATCTCAGCATC	5280
Db	5221	GAGTAAACCTTGGCTGACAGTTTACCAATGCTTAAATCAAGTAGGACCTATCTCAGCATC	5280
QY	5281	TGTCTATTTCGTTATCATCATAGTGGCCGACTCCCGCTCGTATGATATACATACAGTACGG	5340
Db	5281	TGTCTATTTCGTTATCATCATAGTGGCCGACTCCCGCTCGTATGATATACATACAGTACGG	5340
QY	5341	GAGGCGCTTACATCTGGGCCCCAGTGCCTCAATGATATACGCGACAGCCACGCTCACCGCT	5400
Db	5341	GAGGCGCTTACATCTGGGCCCCAGTGCCTCAATGATATACGCGACAGCCACGCTCACCGCT	5400
QY	5401	CCAGATTTTATCAGCAATTAACCCAGCCAGCCGGAAGGGCCGAGCCAGAAAGTGTCTGCA	5460
Db	5401	CCAGATTTTATCAGCAATTAACCCAGCCAGCCGGAAGGGCCGAGCCAGAAAGTGTCTGCA	5460
QY	5461	ACTTTATCCGCTCCATCCATCTATATATTTGTTGGCCGGGAAGCTAGAGTAAGTATTCG	5520
Db	5461	ACTTTATCCGCTCCATCCATCTATATATTTGTTGGCCGGGAAGCTAGAGTAAGTATTCG	5520
QY	5521	CCAGTTATTAATTTGGCGCAAGTGTGGCCATTGTCTACAGCATCTGTGGTCCAGCTCG	5580
Db	5521	CCAGTTATTAATTTGGCGCAAGTGTGGCCATTGTCTACAGCATCTGTGGTCCAGCTCG	5580
QY	5581	TGCTTTGGTATAGCTCTCATTAAGCTCCGCGTTCCCAACGATCAAGCGAGTTTACATGATCC	5640
Db	5581	TGCTTTGGTATAGCTCTCATTAAGCTCCGCGTTCCCAACGATCAAGCGAGTTTACATGATCC	5640
QY	5641	CCCATGTTGTGCAAAAAAGGGGTAGTCCCTGGGTCTCCGATCGTATGTCAGAAATAG	5700
Db	5641	CCCATGTTGTGCAAAAAAGGGGTAGTCCCTGGGTCTCCGATCGTATGTCAGAAATAG	5700
QY	5701	TTGGCCGAGGTTATCACTCATGCTGTTATGCGACACTGTCATTAATTCCTTACAGTCATG	5760
Db	5701	TTGGCCGAGGTTATCACTCATGCTGTTATGCGACACTGTCATTAATTCCTTACAGTCATG	5760
QY	5761	CCATCCGTAAGATGCTTTTCTGTGACGTGGTAGTACTCAACCAAGTCATTTCTGAATAG	5820
Db	5761	CCATCCGTAAGATGCTTTTCTGTGACGTGGTAGTACTCAACCAAGTCATTTCTGAATAG	5820
QY	5821	TGTTATGGGCGACCGAGTTGCTCTGTGCCCGGCGTAAATACGGATATATCCGGCCCAT	5880
Db	5821	TGTTATGGGCGACCGAGTTGCTCTGTGCCCGGCGTAAATACGGATATATCCGGCCCAT	5880
QY	5881	AGCAGAACTTTAAAAAGTGCATCATATTGGAAGAAAGTTCCTTGGGGCGAAAACTCTCAAG	5940
Db	5881	AGCAGAACTTTAAAAAGTGCATCATATTGGAAGAAAGTTCCTTGGGGCGAAAACTCTCAAG	5940
QY	5941	ATCTTACCGGTGTGAGATCCAGTTTCGATGTATACCCACTCGTGCACCCCACTGATCTTCA	6000
Db	5941	ATCTTACCGGTGTGAGATCCAGTTTCGATGTATACCCACTCGTGCACCCCACTGATCTTCA	6000
QY	6001	GCATCTTTACTTTCACCAAGCGTTTCTGGGTGACGCAAAAAACAGAGAAGCAAAATGCGCA	6060
Db	6001	GCATCTTTACTTTCACCAAGCGTTTCTGGGTGACGCAAAAAACAGAGAAGCAAAATGCGCA	6060
QY	6061	AAAAAGGAAATTAAGGGCGACACGGAATGTTGATATCAATACATCTTCCTTTTCAATAT	6120
Db	6061	AAAAAGGAAATTAAGGGCGACACGGAATGTTGATATCAATACATCTTCCTTTTCAATAT	6120
QY	6121	TATTGAACATTTATCAGGGTTATTTGTCTCATGAGCGGATACATATTTGATGATATTAG	6180
Db	6121	TATTGAACATTTATCAGGGTTATTTGTCTCATGAGCGGATACATATTTGATGATATTAG	6180
QY	6181	AAAAATAACAATTAAGGGGTTCCGCGCACATTTCCCGAAAAATGTCAC 6229	

Db	6181	AAAAATAACAAATAGGGGTTCCGCGACATTTCCCGAAAGTGCCAC	6229	Db	6519	TGCTAGAACAGATGCAATGAGATATTAATCACTTTATGGGATCAAAAGCCTAAAGCCATGTG	6578
RESULT 2							
AA096140				Qy	1651	TAAATTAACCCCACTCGTGTAGTTAAAGTGACAGATTGGAAGAAATGACTATA	1770
ID	AA096140	standard; DNA; 9709 BP.		Db	6579	TAAATTAACCCCACTCGTGTAGTTAAAGTGACAGATTGGAAGAAATGACTATA	6638
XX	AA096140:			Qy	1711	CCAAATAGTAGAGCGGAGAAATGATTAATGAGAAAGAGATATAAAATGCTCTTCA	1770
XX	24-FEB-1996	(first entry)		Db	6639	CCAAATAGTAGAGCGGAGAAATGATTAATGAGAAAGAGATATAAAATGCTCTTCA	6638
XX	HIV-1 NL4-3	genomic DNA.		Qy	1771	ATATCAGCAACAGCATAGAGATTAAGGTGAGAAAAGAAATATGATCTTTTATAACTTG	1830
XX	HIV-1:	AIDS; attenuation; vaccine: ss.		Db	6699	ATATCAGCAACAGCATAGAGATTAAGGTGAGAAAAGAAATATGATCTTTTATAACTTG	6758
XX	Human immunodeficiency virus type 1.			Qy	1831	ATATAGTACCAATGATATATACCGCTTACGTTAGTATGATTAAGCCCTCAGCATTA	1890
XX	WO9521912-A1.			Db	6759	ATATAGTACCAATGATATATACCGCTTACGTTAGTATGATTAAGCCCTCAGCATTA	6818
XX	17-AUG-1995.			Qy	1891	CACAGCCCTGTCAAAAGGTATCTTTGAGCCAAATCCCATATATATGTCGCCCGGCTG	1950
XX	14-FEB-1995;	95WO-AU00063.		Db	6819	CACAGCCCTGTCAAAAGGTATCTTTGAGCCAAATCCCATATATATGTCGCCCGGCTG	6878
XX	23-DEC-1994;	94AU-0000284.		Qy	1951	GTTTTGCGATTTCTAAAATGTATATATAAGACGTTCAATGAAACAGACCATGTACAAATG	2010
XX	PR 14-FEB-1994;	94AU-0003864.		Db	6879	GTTTTGCGATTTCTAAAATGTATATATAAGACGTTCAATGAAACAGACCATGTACAAATG	6938
XX	PR 21-FEB-1994;	94AU-0004002.		Qy	2011	TCAGCAGAGTCAATGTACACATGGAATCAGGCCAGTAGTATCAACTCACTGCTGTAA	2070
XX	(AURE-) AUSTRALIAN RED CROSS SOC NSW DIV.			Db	6939	TCAGCAGAGTCAATGTACACATGGAATCAGGCCAGTAGTATCAACTCACTGCTGTAA	6998
XX	(MACF-) MACFARLANE BURNET CENT MEDICAL.			Qy	2071	ATGGCAGCTGTCAGAAAGAGATGATTAATGATGATGTCGCCAAATTTCCAGACAAATGCTA	2130
XX	Cooper D, Crowe S, Deacon NJ, Learmont JC, McPhee DA.			Db	6999	ATGGCAGCTGTCAGAAAGAGATGATTAATGATGATGTCGCCAAATTTCCAGACAAATGCTA	7058
XX	WPI: 1995-293115/38.			Qy	2131	AAACCATATATAGTACAGCTGGAACACATCTGTAGAATTAATGTACAAGACCCAAACA	2190
XX	New non-pathogenic HIV-1 strain carrying a deletion in its nef gene			Db	7059	AAACCATATATAGTACAGCTGGAACACATCTGTAGAATTAATGTACAAGACCCAAACA	7118
XX	or LTR region - can be used in a vaccine to inhibit/reduce			Qy	2191	ATACAGAAAGAAATATCCGTTATCCAGAGGGGACCCAGGAGACATTTGTTACATAGAA	2250
XX	productive infection in an individual by a pathogenic strain			Db	7119	ATACAGAAAGAAATATCCGTTATCCAGAGGGGACCCAGGAGACATTTGTTACATAGAA	7178
XX	Disclosure: Page 8-14; 301pp; English.			Qy	2251	AAATAGGAATATGAGACACACATTTTACATTTAGTAGAGCAAAATGAGATGCCACTT	2310
XX	A genomic nucleotide sequence from pathogenic HIV-1 strain NL4-3			Db	7179	AAATAGGAATATGAGACACACATTTTACATTTAGTAGAGCAAAATGAGATGCCACTT	7238
XX	CC are obt'd. by deletion of portions of the nef gene (AA096141) and/or			Qy	2311	TAAACAGATAGCTAGCAAAATTAAGAGAACATTTGAAATATATAAACAATATCTTTA	2370
XX	LTR region (see AA0961406-Q97166).			Db	7239	TAAACAGATAGCTAGCAAAATTAAGAGAACATTTGAAATATATAAACAATATCTTTA	7298
XX	Sequence 9709 BP; 3421 A; 1756 C; 2366 G; 2166 T; 0 other:			Qy	2371	AGCAATCTCAGAGAGGACCCAGAAATTTGTAAGGACACAGTTTAAATGTCGAGGGGAAT	2430
Query Match	43.0%; Score 2676.2; DB 16; Length 9709;			Db	7299	AGCAATCTCAGAGAGGACCCAGAAATTTGTAAGGACACAGTTTAAATGTCGAGGGGAAT	7358
Best Local Similarity	99.9%; Pred. No. 5.2e-314;			Qy	2431	TTTTCTACTGTAATTTCAACACACAGCTTTTAAATAGTACTGTTTAAATAGTACTTGA	2490
Matches 2678; Conservative	0; Mismatches 3; Indels 0; Gaps 0;			Db	7359	TTTTCTACTGTAATTTCAACACACAGCTTTTAAATAGTACTGTTTAAATAGTACTTGA	7418
Qy	1291	CCATGACAGTGAAGAGAGATATCAGCACTTGTGGAGATGGGGGTGGAATGGGGACCA	1350	Qy	2491	CTGAAGGGTCAAAATTAACCTGAAGAGTGAACACATCACTCCCATGAGAAATTAAC	2550
Db	6219	CAATGAGTGAAGAGAGATATCAGCACTTGTGGAGATGGGGGTGGAATGGGGACCA	6278	Db	7419	CTGAAGGGTCAAAATTAACCTGAAGAGTGAACACATCACTCCCATGAGAAATTAAC	7478
Qy	1351	TGCTCCTGGGATATGATGATCTGTAGTGTCTACAGAAAATTTGGGTCACCGTCTATT	1410	Qy	2551	AATTTATTAACATGTGGGAGAGTAAGAAAAGCAATGTATGCCCCCTCCCATGCTGAC	2610
Db	6279	TGCTCCTGGGATATGATGATCTGTAGTGTCTACAGAAAATTTGGGTCACCGTCTATT	6338	Db	7479	AATTTATTAACATGTGGGAGAGTAAGAAAAGCAATGTATGCCCCCTCCCATGCTGAC	7538
Qy	1411	ATGGGCTACTGTGTGGAAGAGCAACCACTCTATTTTGTGCATCAGATGCTAAAG	1470	Qy	2611	AAATAGATGTTCAATTAATTAATCTGGGCTGCTATTTAACAAGAGATGCTGTAATAACA	2670
Db	6339	ATGGGCTACTGTGTGGAAGAGCAACCACTCTATTTTGTGCATCAGATGCTAAAG	6398	Db	7539	AAATAGATGTTCAATTAATTAATCTGGGCTGCTATTTAACAAGAGATGCTGTAATAACA	7598
Qy	1471	CATATGATACAGAGTACATATATGTTTGGCCACACATGCTGTGTACCAACAGACCCA	1530	Qy	2671	ACAATGGTCCGAGATCTTCAAGACCTGGAAGAGCGGATATGAGGACAAATTTGAGAGATG	2730
Db	6399	CATATGATACAGAGTACATATATGTTTGGCCACACATGCTGTGTACCAACAGACCCA	6458	Db	7599	ACAATGGTCCGAGATCTTCAAGACCTGGAAGAGCGGATATGAGGACAAATTTGAGAGATG	7658
Qy	1531	ACCCACAGAGTATGATTTGGAATGTGACAGAAAATTTTAACATGTGGAATAATGACA	1590				
Db	6459	ACCCACAGAGTATGATTTGGAATGTGACAGAAAATTTTAACATGTGGAATAATGACA	6518				
Qy	1591	TGCTAGAACAGATGATGAGATATTAATCACTTTATGGGATCAAAAGCCTAAAGCCATGTG	1650				

OY	2731	AATTATATAAATATAAAGTACTAAAAAATTGACCATTAAGAGTGTGACCCACCAAGGCCAA	2790
Db	7659	AATTATATAAATATAAAGTACTAAAAAATTGACCATTAAGAGTGTGACCCACCAAGGCCAA	7718
OY	2791	AGAGAGAAGTGTGCAGAGAGAAAAAAGACGATGGGAATAGAGAGCTTTGTCTTGGGT	2850
Db	7719	AGAGAGAAGTGTGCAGAGAGAAAAAAGACGATGGGAATAGAGAGCTTTGTCTTGGGT	7778
OY	2851	TCTTGGGAGCAGCAGGAAGCACTATGGGCTGCACGTCAATGACCTGACGCTACAGGCCA	2910
Db	7779	TCTTGGGAGCAGCAGGAAGCACTATGGGCTGCACGTCAATGACCTGACGCTACAGGCCA	7838
OY	2911	GACAAATATATCTGATATATAGTCGACGACGACGAACAATTTGCTGAGGCTATTGAGCGC	2970
Db	7839	GACAAATATATCTGATATATAGTCGACGACGACGAACAATTTGCTGAGGCTATTGAGCGC	7898
OY	2971	AACAGCATCTGTTCGCACTCAGACTCTGGGGCATCAAAACAGCTCCAGGCAAGAAATCTCG	3030
Db	7899	AACAGCATCTGTTCGCACTCAGACTCTGGGGCATCAAAACAGCTCCAGGCAAGAAATCTCG	7958
OY	3031	CTGAGGAAGAATPACCTAAGAGATACAGAGCTCCTGGGGATTGGGGTTGGCTCTGGAAAC	3090
Db	7959	CTGAGGAAGAATPACCTAAGAGATACAGAGCTCCTGGGGATTGGGGTTGGCTCTGGAAAC	8018
OY	3091	TCATTTGCACACATGCTGTGCTGCTTGGAAATGCTAGTTGGATTAATTAATCTCTGGACAGA	3150
Db	8019	TCATTTGCACACATGCTGTGCTGCTTGGAAATGCTAGTTGGATTAATTAATCTCTGGACAGA	8078
OY	3151	TTTGGAAATPAACATGACCTGATGGAGTGGAGGACAGAGAAATPAACAAATPACACAGCTTAA	3210
Db	8079	TTTGGAAATPAACATGACCTGATGGAGTGGAGGACAGAGAAATPAACAAATPACACAGCTTAA	8138
OY	3211	TACACCTCCTTAATTGAAGAAATCGCAAAACACGACAAAGAAATGAACAAATTAATTG	3270
Db	8139	TACACCTCCTTAATTGAAGAAATCGCAAAACACGACAAAGAAATGAACAAATTAATTG	8198
OY	3271	AATTAGATTAATATGGGCAAGTTTGTGGAAATTTGGTTAACTATACAAATTTGGCTGTGTATA	3330
Db	8199	AATTAGATTAATATGGGCAAGTTTGTGGAAATTTGGTTAACTATACAAATTTGGCTGTGTATA	8258
OY	3331	TAAATATTTTCATATATGATAGTATAGAGAGCTTGTAAGTATTAAGATATGTTTTCCTGTAC	3390
Db	8259	TAAATATTTTCATATATGATAGTATAGAGAGCTTGTAAGTATTAAGATATGTTTTCCTGTAC	8318
OY	3391	TTTCTATATGTGAATAGAGTATAGCAGGGATATTCACCATTAATTCGTTTCAGACCCACCTCC	3450
Db	8319	TTTCTATATGTGAATAGAGTATAGCAGGGATATTCACCATTAATTCGTTTCAGACCCACCTCC	8378
OY	3451	CAATCCGAGGGAGCCGACAGGGCCGGAAGGAATAGAAAGAAAGAGTGGAGAGACACA	3510
Db	8379	CAATCCGAGGGAGCCGACAGGGCCGGAAGGAATAGAAAGAAAGAGTGGAGAGACACA	8438
OY	3511	GAGACAGATCCATTGATTTAGTGAAGGATCCTTAGCACTTAATCTGGAGACATTTGCGGA	3570
Db	8439	GAGACAGATCCATTGATTTAGTGAAGGATCCTTAGCACTTAATCTGGAGACATTTGCGGA	8498
OY	3571	GCCCTGTGCTTTTACGACTACACCGGCTTGAGAGACTTACTTGTATTTGTAACAGAGATTG	3630
Db	8499	GCCCTGTGCTTTTACGACTACACCGGCTTGAGAGACTTACTTGTATTTGTAACAGAGATTG	8558
OY	3631	TGGAACCTTCTGGAGCGCAGGGGGTGGGAAGCCCTCAATATATTTGGTGAAATCCTACAGT	3690
Db	8559	TGGAACCTTCTGGAGCGCAGGGGGTGGGAAGCCCTCAATATATTTGGTGAAATCCTACAGT	8618
OY	3691	ATTGGAAGTCAGGAACATAAAGAAATAGTGTCTTTAACTTGTCTCAATGCCACAGCCATPACAG	3750
Db	8619	ATTGGAAGTCAGGAACATAAAGAAATAGTGTCTTTAACTTGTCTCAATGCCACAGCCATPACAG	8678
OY	3751	TAGCTGAGGGGACAGATATAGGTTTATAGAAATATTACAAGACGCTTATATAGACTATTGCC	3810
Db	8679	TAGCTGAGGGGACAGATATAGGTTTATAGAAATATTACAAGACGCTTATATAGACTATTGCC	8738

Query Match	43.0%	Score 2676.2	DB 20	Length 9709
Best Local Similarity	99.9%	Pred. No. 5.2e-314		
Matches 2678	Conservative 0	Mismatches 3	Indels 0	Gaps 0
OY	1291	CCATGAGAGTAAAGAGAAATGATCAGCACTTGTGGAGATGGGGGTGAATAATGGGGCACCA	1350	
DB	6219	CAATGAGAGTAAAGAGAAATGATCAGCACTTGTGGAGATGGGGGTGAATAATGGGGCACCA	6278	
OY	1351	TGCTCCTGGGATATTTGATGATCTGTAGTGTCTACAGAAAATTTGGGTCAACCGTCTATT	1410	
DB	6279	TGCTCCTGGGATATTTGATGATCTGTAGTGTCTACAGAAAATTTGGGTCAACAGTCTATT	6338	
OY	1411	ATGGAGGATCTGTGTGGAAGGACCAACCACTCTATTATTTGTGCATCAGATGCTAAAG	1470	
DB	6339	ATGGAGGATCTGTGTGGAAGGACCAACCACTCTATTATTTGTGCATCAGATGCTAAAG	6398	

1471 CATATGATACAGAGGTACATATGTTGGGCCACACATGCGTGTACCCACAGACCCCA 1530
6399 CATATGATACAGAGGTACATATGTTGGGCCACACATGCGTGTACCCACAGACCCCA 6458
1531 ACCCACAAGAGTATGTTGGTAAATGTCAGAGAAAATTTTAACTGTGGAAAAATGACA 1590
6459 ACCCACAAGAGTATGTTGGTAAATGTCAGAGAAAATTTTAACTGTGGAAAAATGACA 6518
1551 TGTGTAAACAGATGATGAGATATATCACTTTATGGGATCAAAAGCCTAAAGCCATGTG 1650
6519 TGTGTAAACAGATGATGAGATATATCACTTTATGGGATCAAAAGCCTAAAGCCATGTG 6578
1651 TAAATTAACCCACTGCTGTGTAGTTAAAGTGCACATGTTGAAGATATCTATA 1710
6579 TAAATTAACCCACTGCTGTGTAGTTAAAGTGCACATGTTGAAGATATCTATA 1710
1711 CCAATAGTAGAGGGGAGATGATATGAGAAAAGAGATATAAAACTGCTTTCA 1770
6639 CCAATAGTAGAGGGGAGATGATATGAGAAAAGAGATATAAAACTGCTTTCA 6698
1771 ATATCAGCACAAGCATTAAGATTAAGTGCAGAAAAGATATGATCTTTTATAAACTTG 1830
6699 ATATCAGCACAAGCATTAAGATTAAGTGCAGAAAAGATATGATCTTTTATAAACTTG 6758
1831 ATATAGTACCAATAGATATATACAGCTATAGTTGATTAAGTTTAACTGCTGATTA 1890
6759 ATATAGTACCAATAGATATATACAGCTATAGTTGATTAAGTTTAACTGCTGATTA 6818
1891 CACAGGCGCTGTCGCAAGGTATCTTTGAGCCAAATCCCATATATGTCGCCGCGTG 1950
6819 CACAGGCGCTGTCGCAAGGTATCTTTGAGCCAAATCCCATATATGTCGCCGCGTG 6878
1951 GTTTTGGCATCTTAAATGTAATTAAGACGTTCAATGAGAACAGGACCATGTAATG 2010
6879 GTTTTGGCATCTTAAATGTAATTAAGACGTTCAATGAGAACAGGACCATGTAATG 6938
2011 TCAGACAGTACATGTACATGGAATCAGGCCAGTGTATCAACTCACTGCTGTTAA 2070
6939 TCAGACAGTACATGTACATGGAATCAGGCCAGTGTATCAACTCACTGCTGTTAA 6998
2071 ATGGCAGTCTAGCAGAAAGATGTAGTAATTAATGTCGCCAATTTTCACAGACAATGCTA 2130
6999 ATGGCAGTCTAGCAGAAAGATGTAGTAATTAATGTCGCCAATTTTCACAGACAATGCTA 7058
2131 AAACCATTAATGTACAGCTGACACATCTGTAGAATTAATTTGTAACAAGACCCAAACA 2190
7059 AAACCATTAATGTACAGCTGACACATCTGTAGAATTAATTTGTAACAAGACCCAAACA 7118
2191 ATACAGAAAAAGTATCCGTATCCAGAGGGGACAGGAGGACATTTGTTCAATAGGAA 2250
7119 ATACAGAAAAAGTATCCGTATCCAGAGGGGACAGGAGGACATTTGTTCAATAGGAA 7178
2251 AAATAGGAATATGAGACAAACACATTTAGTAGAGCAAAATGATGCCACTT 2310
7179 AAATAGGAATATGAGACAAACACATTTAGTAGAGCAAAATGATGCCACTT 7238
2311 TAAACAGATAGCTAGCAAAATTAAGAGAACATTTGCAATTAATTAATTAATCTTTA 2370
7239 TAAACAGATAGCTAGCAAAATTAAGAGAACATTTGCAATTAATTAATTAATCTTTA 7298
2371 AGCATCTCTAGAGGGGACCCAGAAATTTGTAAGGACAGTTTAAATGTTGAGGGGAT 2430
7299 AGCATCTCTAGAGGGGACCCAGAAATTTGTAAGGACAGTTTAAATGTTGAGGGGAT 7358
2431 TTTTCTACTGTATTAACACAACTGTTTAAATAGTACTTGTAAATAGTACTTGTGA 2490
7359 TTTTCTACTGTATTAACACAACTGTTTAAATAGTACTTGTAAATAGTACTTGTGA 7418
2491 CTGAAGGCTAAATTAACACTGTAAGAGAGTGCACAACTCCCATGCAAGATTA 2550
7419 CTGAAGGCTAAATTAACACTGTAAGAGAGTGCACAACTCCCATGCAAGATTA 7478

2551 AATTTAATACATGTCGAGGAGTAGGAAAAAGCAATGATCCCTCCATCAGTGGAC 2610
7479 AATTTAATACATGTCGAGGAGTAGGAAAAAGCAATGATCCCTCCATCAGTGGAC 7538
2611 AAATAGATGTTTCAATTAATATTTACTGGGCTGCTATTAACAGAGATGTTGTAATACA 2670
7539 AAATAGATGTTTCAATTAATATTTACTGGGCTGCTATTAACAGAGATGTTGTAATACA 7598
2671 ACAATGGTCCGAGATCTTACAGACCTGAGAGAGCGGATATAGGAGCAATTTGAGAGCTG 2730
7599 ACAATGGTCCGAGATCTTACAGACCTGAGAGAGCGGATATAGGAGCAATTTGAGAGCTG 7658
2731 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2790
7659 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7718
2791 AGAGAAGAGTGTGTCAGAGAAAAAAGAGAGTGGGAATAGAGACTTGTCTTGCGT 2850
7719 AGAGAAGAGTGTGTCAGAGAAAAAAGAGAGTGGGAATAGAGACTTGTCTTGCGT 7778
2851 TCTTGGGAGCAGAGGAAGCACTATGGGCTGCACATGACGCTGACGGTACAGGCCA 2910
7779 TCTTGGGAGCAGAGGAAGCACTATGGGCTGCACATGACGCTGACGGTACAGGCCA 7838
2911 GACAATTAATGCTGATTAATGTCAGAGCAAGAACAAATTTGCTGAGGCTATTTGAGCGC 2970
7839 GACAATTAATGCTGATTAATGTCAGAGCAAGAACAAATTTGCTGAGGCTATTTGAGCGC 7898
2971 AACAGCATCTGTTCAACTCACAGCTGAGGAGATCAAAACACTCCAGCAGAAATCTCGG 3030
7899 AACAGCATCTGTTCAACTCACAGCTGAGGAGATCAAAACACTCCAGCAGAAATCTCGG 7958
3031 CTGTGGAAGATTAACCTTAAGATCAACAGCTCTGAGGATTTGGGGTCTCTGGA 3090
7959 CTGTGGAAGATTAACCTTAAGATCAACAGCTCTGAGGATTTGGGGTCTCTGGA 8018
3091 TCATTTGACACAGCTGCTGCTGGAATGCTAGTGGAGTAATTAATCTCTGGAACGA 3150
8019 TCATTTGACACAGCTGCTGCTGGAATGCTAGTGGAGTAATTAATCTCTGGAACGA 8078
3151 TTTGGAATTAACATGACCTGATGAGTGGAGTGGACAGAGAAATTAACAAATTCACAAGCTTAA 3210
8079 TTTGGAATTAACATGACCTGATGAGTGGAGTGGACAGAGAAATTAACAAATTCACAAGCTTAA 8138
3211 TACACTGCTTAATTTGAAGATTCGCAAAAACGCAAGAAAGATTAACAAATTTATGG 3270
8139 TACACTGCTTAATTTGAAGATTCGCAAAAACGCAAGAAAGATTAACAAATTTATGG 8198
3271 AATTAGATTAATTTGGCAAGTTTGGGAATGTTTAACATTAACAAATTTGGCTGAGTATA 3330
8199 AATTAGATTAATTTGGCAAGTTTGGGAATGTTTAACATTAACAAATTTGGCTGAGTATA 8258
3331 TAAATTAATTCATATATGATAGTACAGGCTTGTGATAGTTTAAAGATAGTTTCTGTAC 3390
8259 TAAATTAATTCATATATGATAGTACAGGCTTGTGATAGTTTAAAGATAGTTTCTGTAC 8318
3391 TTTCTATAGTAATAGATTAAGCAGGATATTTACCATTAATGCTTTTACAGCCACTCC 3450
8319 TTTCTATAGTAATAGATTAAGCAGGATATTTACCATTAATGCTTTTACAGCCACTCC 8378
3451 CAATCCGAGGGGACCCGACAGGCGCGAAGGAATGAAGAAAGAGGGAGAGAGACA 3510
8379 CAATCCGAGGGGACCCGACAGGCGCGAAGGAATGAAGAAAGAGGGAGAGAGACA 8438
3511 GAGACAGATCCATTCGATTAATGTAAGCGATCTTATGACACTATCTGGAGAGATCTGGGGA 3570
8439 GAGACAGATCCATTCGATTAATGTAAGCGATCTTATGACACTATCTGGAGAGATCTGGGGA 8498
3571 GCTGTGCTCTTATAGCTACACGCGTTTGAGAGACTTAATCTTATTTGAACAGGATTTG 3630
8499 GCTGTGCTCTTATAGCTACACGCGTTTGAGAGACTTAATCTTATTTGAACAGGATTTG 8558
3631 TGAATCTTCTGAGCGCAGGGGGTGGGAAGCCCTCAATATATGTTGGAATCTCTACAGT 3690

|||||
Db 8559 TGGAACTTCTGGGACGAGGGGGTGGGAACCCCTCAATATGTGTGAATCTCTACAGT 8618
Qy 3691 ATGGAGTCGAGAACTAAAGATAGTCTGTACCTTGCATTCGCCACGACATGACAG 3750
Db 8619 ATTTGAGTCGAGAACTAAAGATAGTCTGTACCTTGCATTCGCCACGACATGACAG 8678
Qy 3751 TACGTAGGGGACAGATAGGGTTATAGAACTATTACAGCAGCTTATAGACTATTCGCC 3810
Db 8679 TACTGAGGGGACAGATAGGGTTATAGAACTATTACAGCAGCTTATAGACTATTCGCC 8738
Qy 3811 ACATACCTAGAAATTAAGACAGCGCTTGGAAAAGATTTTCTATTAAGATGGGTGCAAG 3870
Db 8739 ACATACCTAGAAATTAAGACAGCGCTTGGAAAAGATTTTCTATTAAGATGGGTGCAAG 8798
Qy 3871 TGGTCAAAAAGTATGTATGTATGGCTGTCTGTATAGGAAAGATATGAGTGGTGGCAAG 3930
Db 8799 TGGTCAAAAAGTATGTATGTATGGCTGTCTGTATAGGAAAGATATGAGTGGTGGCAAG 8858
Qy 3931 CCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAGATCTAGA 3971
Db 8859 CCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAGACCTAGA 8899

RESULT 4

AAA97926 standard; DNA: 9709 BP.

AAA97926;

26-JAN-2001 (first entry)

HIV-1 env DNA.

Viral protein: vaccine; anti-viral; anti-HIV; therapy; infection;

env protein; envelope protein; ds.

Human immunodeficiency virus type 1.

W0200047223-A2.

17-AUG-2000.

03-DEC-1999; 99WO-EP09759.

12-FEB-1999; 99DE-1007485.

(STRA-) STRATHMANN & CO AG.

Schreiber M;

WPI: 2000-549084/50.

P-PSDB: AAB10697.

PS Claim 11: Page 68-70; 79pp; German.
Viral vaccine comprises a mixture of protein sequence variants of a single viral protein, which is useful for prevention and therapy of viral infections, especially HIV, in humans

This invention describes a novel protein vaccine which comprises a mixture of viral proteins, characterized in that the molecules are sequence variants of a single viral (partial) protein. The products of the invention have anti-viral and anti-HIV activity. Mixtures of structurally different viral proteins, that are sequence variants of a single protein are useful for production of vaccines for the prevention and/or therapy of viral infections in humans. The vaccines are especially useful for prevention and/or therapy of human immunodeficiency virus (HIV) infection in humans. The vectors and host cells are useful for the expression of the protein/DNA mixtures, which are also useful for preventing and/or therapy of viral infection. This sequence encodes the human immunodeficiency virus (HIV-1) envelope (env) protein which is used to illustrate the method of the invention.

xx
SQ Sequence 9709 BP; 3421 A; 1756 C; 2366 G; 2166 T; 0 other;
Query Match 43.0%; Score 2676.2; DB 21; Length 9709;
Best Local Similarity 99.9%; Pred. No. 5.2e-314;
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1291 CCATGAGAGTGAAGAGAAAGTATCACCACCTTGTGAGATGGGGTGAATGAGGACCA 1350
Db 6219 CAATGAGAGTGAAGAGAAAGTATCACCACCTTGTGAGATGGGGTGAATGAGGACCA 6278
Qy 1351 TGGTCTTGGGATTTGATGATCTGTAGTCTACAGAAAATTTGGTGCCTTATT 1410
Db 6279 TGGTCTTGGGATTTGATGATCTGTAGTCTACAGAAAATTTGGTGCCTTATT 6338
Qy 1411 ATGGGGTACTGTGTGGAAAGCAACCCACTCTATTTTGTGCATCGATGCTAAAG 1470
Db 6339 ATGGGGTACTGTGTGGAAAGCAACCCACTCTATTTTGTGCATCGATGCTAAAG 6398
Qy 1471 CATATGATACAGAGTACATATATGTTTGGGCCACAGATGCCCTGTACCCAGACCCCA 1530
Db 6399 CATATGATACAGAGTACATATATGTTTGGGCCACAGATGCCCTGTACCCAGACCCCA 6458
Qy 1531 ACCCACAAGAGTATGTTTGGTTAAAGTGCACAAAAATTTTAACTGTGGAAAAATGACA 1590
Db 6459 ACCCACAAGAGTATGTTTGGTTAAAGTGCACAAAAATTTTAACTGTGGAAAAATGACA 6518
Qy 1591 TGGTAGACAGATGATGAGATATATATGTTTAAAGTGCACAAAAATTTTAACTGTGG 1650
Db 6519 TGGTAGACAGATGATGAGATATATATGTTTAAAGTGCACAAAAATTTTAACTGTGG 6578
Qy 1651 TAAATTTAACCCACTCTGTGTAAAGTGCACAAAAATTTTAACTGTGGAAAAATGACA 1710
Db 6579 TAAATTTAACCCACTCTGTGTAAAGTGCACAAAAATTTTAACTGTGGAAAAATGACA 6638
Qy 1711 CCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1770
Db 6639 CCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6698
Qy 1771 ATATGACGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1830
Db 6699 ATATGACGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6758
Qy 1831 ATATGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1890
Db 6759 ATATGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6818
Qy 1891 CACAGGCTGTCCAAAGGTATCCTTTGAGCCATTTCCCATACATATTATTTGCCCGGCTG 1950
Db 6819 CACAGGCTGTCCAAAGGTATCCTTTGAGCCATTTCCCATACATATTATTTGCCCGGCTG 6878
Qy 1951 GTTTTGGGATTTTAAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2010
Db 6879 GTTTTGGGATTTTAAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6938
Qy 2011 TCAGCAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2070
Db 6939 TCAGCAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6998
Qy 2071 ATGGCAGTACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2130
Db 6999 ATGGCAGTACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7058
Qy 2131 AAACGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2190
Db 7059 AAACGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7118
Qy 2191 ATACAAAGAAAGATATCCGATATCCAGAGGGGACAGGAGAGAGATTTGTTTACAAATGAGAA 2250
Db 7119 ATACAAAGAAAGATATCCGATATCCAGAGGGGACAGGAGAGAGATTTGTTTACAAATGAGAA 7178
Qy 2251 AAATAGGAATATGAGACAAAGCATTGTACATTTAGATGAGCAAAATGGAATGCCACTT 2310
|||||

QY	3391	TTTCTATAGTAATAGAGTTAGGACGGGATATTCCACATTATTCGTTTCAGACCCACCTCC	3450
Db	8319	TTTCTATAGTAATAGAGTTAGGACGGGATATTCCACATTATTCGTTTCAGACCCACCTCC	8378
QY	3451	CAATCCGAGGGGACCCGACAGGCCCCGAGGAAATAGAAAGGAGTGAGAGAGACACA	3510
Db	8379	CAATCCGAGGGGACCCGACAGGCCCCGAGGAAATAGAAAGGAGTGAGAGAGAGACACA	8438
QY	3511	GAGACAGATTCATTTCGATTAGTAGAGGATCCTTACGACTATATCTGGAGCATGTGGGA	3570
Db	8439	GAGACAGATTCATTTCGATTAGTAGAGGATCCTTACGACTATATCTGGAGCATGTGGGA	8498
QY	3571	GCCTGTGCTTCCTACGCTACCCAGCCGCTTGAGAGACTTACTCTTGATTGTAAAGAGATTG	3630
Db	8499	GCCTGTGCTTCCTACGCTACCCAGCCGCTTGAGAGACTTACTCTTGATTGTAAAGAGATTG	8558
QY	3631	TGGAACCTTCCTGGAGCGAGGGGGGTGGGAAGCCCAATATATGGTGGAAATCCTACAGT	3690
Db	8559	TGGAACCTTCCTGGAGCGAGGGGGGTGGGAAGCCCAATATATGGTGGAAATCCTACAGT	8618
QY	3691	ATTGGAGTCAGGAACCTAAAGAAATAGTGTCTTAACTTGCTCAATGCCACAGCCATPACAG	3750
Db	8619	ATTGGAGTCAGGAACCTAAAGAAATAGTGTGTCTTAACTTGCTCAATGCCACAGCCATPACAG	8678
QY	3751	TAGCTGAGGGGACAGATAGGGTTTATAGAAATATACAAAGCCTTATAGAGCTATTGGCC	3810
Db	8679	TAGCTGAGGGGACAGATAGGGTTTATAGAAATATACAAAGCCTTATAGAGCTATTGGCC	8738
QY	3811	ACATACCTAGAAAGATTAAGACAGGGCTTGGAAAGATTTTTCCTAATAGATGGGTGGCAAG	3870
Db	8739	ACATACCTAGAAAGATTAAGACAGGGCTTGGAAAGATTTTTCCTAATAGATGGGTGGCAAG	8798
QY	3871	TGGGTCAAAAAGTACTGTGATTTGGATTTGGCTGTCTAGAGGAAAGAAATGAGACGAGCTGAG	3930
Db	8799	TGGGTCAAAAAGTACTGTGATTTGGATTTGGCTGTCTAGAGGAAAGAAATGAGACGAGCTGAG	8858
QY	3931	CCAGCAGCAGATGGGGTGGAGCAGTATCTTCGAGATCTAGA 3971	
Db	8859	CCAGCAGCAGATGGGGTGGAGCAGTATCTTCGAGATCTAGA 8899	
RESULT 5			
AAAA0298			
Id	AAAA0298	standard; DNA; 9709 BP.	
XX	AC	AAAA0298;	
XX	AC		
DT	02-NOV-2000	(first entry)	
XX	XX		
DE	HIV-1 viral protein operon encoding polynucleotide.		
XX	XX		
KW	Glycoprotein; gag gene; pol gene; GP-1; GP-2; anti-HIV; cytostatic;		
KW	gene therapy; treatment; infectious disease; HIV; AIDS; neoplasm;		
KW	carcinoma; melanoma; env gene; viral protein; ds.		
OS	Human immunodeficiency virus type 1.		
XX	XX		
PN	Ep1006196-A2.		
XX	XX		
PD	07-JUN-2000.		
XX	XX		
PF	25-NOV-1999;	99EP-0250415.	
XX	XX		
PR	26-NOV-1998;	98DE-1056463.	
XX	XX		
PA	(PERT-) PETTE INST HEINRICH.		
XX	XX		
PI	Von Laer MD;		
XX	XX		
DR	WP1; 2000-378268/33.		
DR	P-PSDB; AAB10046; AAB10047; AAB10048; AAB10049; AAB10050; AAB10051.		
DR	AAB10052; AAB10053; AAB10054.		

XX New retroviral packing cell useful as pharmaceutical carrier in gene
PT therapy for treatment of HIV and neoplasms, comprises retroviral genes
PT and glycoproteins -

PS Disclosure: Page 37-40; 69pp; German.

XX
XX This invention describes a novel retroviral packing cell (1), comprising
CC the retroviral genes gag, pol and glycoproteins gp-1 and gp-2 of the LCMV
CC coding gene gp, or a part of these. The products of the invention have
CC anti-HIV and cytostatic activity and can be used for gene therapy. (1) is
CC useful for in vitro infection of cells, especially hematopoietic stem
CC cells, for expression of transgenes in cells and as a pharmaceutical
CC carrier for gene therapy. (1) is therefore useful in the treatment of
CC infectious diseases (HIV or AIDS), neoplasms (carcinomas), melanomas and
CC other diseases. This sequence encodes the Human immunodeficiency virus
CC (HIV-1) gag, pol, vif, vpr, tat, rev, vpu, coat and nef proteins which
CC are described in the method of the invention.

XX Sequence 9709 BP; 3421 A; 1756 C; 2366 G; 2166 T; 0 other;

Query Match 43.0%; Score 2676.2; DB 21; Length 9709;

Best Local Similarity 99.9%; Pred. No. 5.2e-314;

Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1291	CCATGAGAGTGAAGAGAGATATCAGCCTTGTGAGATGGGGGTGGAATGGGGCCACA	1350
DB	6219	CAATGAGAGTGAAGAGAGATATCAGCCTTGTGAGATGGGGGTGGAATGGGGCCACA	6278
QY	1351	TGCTCCCTGGGATTTATGATATCTGTAGTGTCTACAGAAAATTTGGTCCCTCTATT	1410
DB	6379	TGCTCCCTGGGATTTATGATATCTGTAGTGTCTACAGAAAATTTGGTCCCTCTATT	6338
QY	1411	ATGGGGTACCTGTGTGAGAGAGACACCACTCTATTGTGATCAGATCTTAAG	1470
DB	6339	ATGGGGTACCTGTGTGAGAGAGACACCACTCTATTGTGATCAGATCTTAAG	6388
QY	1471	CATATGATACAGAGTACATTAATGTTGGCCACACATGCCGTGTACCCACAGACCCA	1530
DB	6399	CATATGATACAGAGTACATTAATGTTGGCCACACATGCCGTGTACCCACAGACCCA	6458
QY	1531	ACCCACAAGAGTGTATGTTGTAATGTGACAGAAAATTTTACATGTGGAATAATGACA	1590
DB	6459	ACCCACAAGAGTGTATGTTGTAATGTGACAGAAAATTTTACATGTGGAATAATGACA	6518
QY	1591	TGCTAGAACAGATCATGAGGATATATCATGTTATGGATCAAGCCTTAAGCCATGTG	1650
DB	6519	TGCTAGAACAGATCATGAGGATATATCATGTTATGGATCAAGCCTTAAGCCATGTG	6578
QY	1651	TAAATTTAACCCACTCTGTGTATGTTTAAAGTGCACATGTTTGAAGATGATACTATA	1710
DB	6579	TAAATTTAACCCACTCTGTGTATGTTTAAAGTGCACATGTTTGAAGATGATACTATA	6638
QY	1711	CCATATAGTAGAGGGAGATGATATGAGAAAAGAGATTAATAAACTGCTCTTCA	1770
DB	6639	CCATATAGTAGAGGGAGATGATATGAGAAAAGAGATTAATAAACTGCTCTTCA	6698
QY	1771	ATATCAGACAGCATTAAGATTAAGTGCAGAAAAGATTAATGCTCTTTTAACTTG	1830
DB	6699	ATATCAGACAGCATTAAGATTAAGTGCAGAAAAGATTAATGCTCTTTTAACTTG	6758
QY	1831	ATATAGTACCAATTAATTAATACAGCTATAGTTGATTAAGTTTACACCTCACTCATTA	1890
DB	6759	ATATAGTACCAATTAATTAATACAGCTATAGTTGATTAAGTTTACACCTCACTCATTA	6818
QY	1891	CACAGGCTCTCCAAAGTATCCTTTGAGCCAATCCCATACATATTATTGCCCCGCTG	1950
DB	6819	CACAGGCTCTCCAAAGTATCCTTTGAGCCAATCCCATACATATTATTGCCCCGCTG	6878
QY	1951	GTTTGGGATTTCTAAATGTAATTAAGAGCTTCATGGAACGAGCATGTAACAATG	2010
DB	6879	GTTTGGGATTTCTAAATGTAATTAAGAGCTTCATGGAACGAGCATGTAACAATG	6938

QY	2011	TCAGCAGTACAAATGATACATGAAATCAGGCCAGTAGATATCACTCACTGCTTTAA	2070
DB	6939	TCAGCAGTACAAATGATACATGAAATCAGGCCAGTAGATATCACTCACTGCTTTAA	6998
QY	2071	ATGGCAGTCTAGCAGAAAGATGTAGTAATTAATCTGCCAATTTTACAGACAAATGCTA	2130
DB	6999	ATGGCAGTCTAGCAGAAAGATGTAGTAATTAATCTGCCAATTTTACAGACAAATGCTA	7058
QY	2131	AAACCATTAATAGTACAGTCAACATCTGTAGAAATTAATTTGACAAAGCCCAACACA	2190
DB	7059	AAACCATTAATAGTACAGTCAACATCTGTAGAAATTAATTTGACAAAGCCCAACACA	7118
QY	2191	ATACAGAAAAAGTATCCGATATCCAGAGGGACCCAGGAGACATTTGTTACAAATAGAA	2250
DB	7119	ATACAGAAAAAGTATCCGATATCCAGAGGGACCCAGGAGACATTTGTTACAAATAGAA	7178
QY	2251	AAATAGGAATATATAGACAAACATTTGTAACTTATAGAGCAAAATGGAAATGCCACTT	2310
DB	7179	AAATAGGAATATATAGACAAACATTTGTAACTTATAGAGCAAAATGGAAATGCCACTT	7238
QY	2311	TAAACAGATATAGTACCAATTAAGAGACAAATTTGAAATTAATTAATTAATCTTTA	2370
DB	7239	TAAACAGATATAGTACCAATTAAGAGACAAATTTGAAATTAATTAATTAATCTTTA	7298
QY	2371	AGCAATCCTCAGAGGGGACCCAGAAATTTGTAACGACAGTTTAAATTTGGAGGGAAT	2430
DB	7299	AGCAATCCTCAGAGGGGACCCAGAAATTTGTAACGACAGTTTAAATTTGGAGGGAAT	7358
QY	2431	TTTTCTACTGTATATTCACACAACTGTTTAAATTAATCTGTTTAAATAGTCTTGAGTA	2490
DB	7359	TTTTCTACTGTATATTCACACAACTGTTTAAATTAATCTGTTTAAATAGTCTTGAGTA	7418
QY	2491	CTGAAGGGTCAAAATTAACAGTAAGAGAGTGACACAACTCCATGAGAAATTAAC	2550
DB	7419	CTGAAGGGTCAAAATTAACAGTAAGAGAGTGACACAACTCCATGAGAAATTAAC	7478
QY	2551	AATTTATTAACATGTGCGAGAAAGTGAAGAAAAGCAATGTATGCCCTCCATCAGTGAC	2610
DB	7479	AATTTATTAACATGTGCGAGAAAGTGAAGAAAAGCAATGTATGCCCTCCATCAGTGAC	7538
QY	2611	AAATTAATGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2670
DB	7539	AAATTAATGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	7598
QY	2671	ACAATGCGTCCGAGATTTTCAGACTTGCAGAGAGGCGATATGAGGACAAATTTGAGAA	2730
DB	7599	ACAATGCGTCCGAGATTTTCAGACTTGCAGAGAGGCGATATGAGGACAAATTTGAGAA	7658
QY	2731	AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2790
DB	7659	AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	7718
QY	2791	AGAGAAAGTGTGTCAGAGAGAAAAGACAGTGGGGAATTAAGAGCTTTGCTCTGGGT	2850
DB	7719	AGAGAAAGTGTGTCAGAGAGAAAAGACAGTGGGGAATTAAGAGCTTTGCTCTGGGT	7778
QY	2851	TCTTGGAGCAGCAGAGAAAGCCTATGCGCTGCAAGTCAATGACGCTGACGAGGCCA	2910
DB	7779	TCTTGGAGCAGCAGAGAAAGCCTATGCGCTGCAAGTCAATGACGCTGACGAGGCCA	7838
QY	2911	GACATTAATTTGCTGATATATAGTGCAGAGAGAAACAAATTTGCTGAGGGCTATTTAG	2970
DB	7839	GACATTAATTTGCTGATATATAGTGCAGAGAGAAACAAATTTGCTGAGGGCTATTTAG	7898
QY	2971	AACAGATCTGTTGCAACATCACTGCGGGGCAATCAACAGCTCCAGGCAAGAAATCCTG	3030
DB	7899	AACAGATCTGTTGCAACATCACTGCGGGGCAATCAACAGCTCCAGGCAAGAAATCCTG	7958
QY	3031	CTGTGAGAAAGATTAATTAAGGATCAACAGCTCTGCGGGATTTGGGGTGTCTGGA	3090
DB	7959	CTGTGAGAAAGATTAATTAAGGATCAACAGCTCTGCGGGATTTGGGGTGTCTGGA	8018
QY	3091	TCAATTTGACCACTGCTGCTGCTGGAATGCTAGTTGGAATTAATTAATCTCTGGA	3150

```

Db      8019 TCATTTGACCACTCTGCTGCTTGAGTGTGAGTATGATTAATCTCTGGACACAGA 8078
      |||
Qy      3151 TTGGCAATACATGACCTGTGATGAGTGGGACAGAGAAATTAACAATTAACACAGCTTAA 3210
      |||
Db      8079 TTGGCAATACATGACCTGTGATGAGTGGGACAGAGAAATTAACAATTAACACAGCTTAA 8138
      |||
Qy      3211 TACACTCCTTAATTAAGAAATTCGCAAAACCAAGAAAGAAATTAACAATTAATGG 3270
      |||
Db      8139 TACACTCCTTAATTAAGAAATTCGCAAAACCAAGAAAGAAATTAACAATTAATGG 8198
      |||
Qy      3271 AATTGATTAATGGGCAAGTTTGTGAAATTTGTTTAACATTAACAATTTGGCTGTGATA 3330
      |||
Db      8199 AATTGATTAATGGGCAAGTTTGTGAAATTTGTTTAACATTAACAATTTGGCTGTGATA 8258
      |||
Qy      3331 TAAATTAATTAATGATAGTAGAGGCTTGTGATTAAGAAATTTTGGCTGTAC 3390
      |||
Db      8259 TAAATTAATTAATGATAGTAGAGGCTTGTGATTAAGAAATTTTGGCTGTAC 8318
      |||
Qy      3391 TTTGATAGTGAATAGATTAGGCGAGGATATTCACCATTTATCCTTTAGACCCCTTCC 3450
      |||
Db      8319 TTTGATAGTGAATAGATTAGGCGAGGATATTCACCATTTATCCTTTAGACCCCTTCC 8378
      |||
Qy      3451 CAATCCGAGGGGACCCGACAGGCCCGAAGAAATAGAGAAAGGTGAGAGAGACAA 3510
      |||
Db      8379 CAATCCGAGGGGACCCGACAGGCCCGAAGAAATAGAGAAAGGTGAGAGAGACAA 8438
      |||
Qy      3511 GAGACAGATTCATTCGATTAGTGAAGCGATCTTATGACATTAATCTGAGAGATCTGCGGA 3570
      |||
Db      8439 GAGACAGATTCATTCGATTAGTGAAGCGATCTTATGACATTAATCTGAGAGATCTGCGGA 8498
      |||
Qy      3571 GCCTGTGCTCTTCAGTACACCCGCTTGAGAGATTTCTGATTTGTAACGAGAGATTG 3630
      |||
Db      8499 GCCTGTGCTCTTCAGTACACCCGCTTGAGAGATTTCTGATTTGTAACGAGAGATTG 8558
      |||
Qy      3631 TGGAACTTCTGGAGACGAGGGGTGGGAAGCCCTCAATATTTGTTGGAATCTCTACAGT 3690
      |||
Db      8559 TGGAACTTCTGGAGACGAGGGGTGGGAAGCCCTCAATATTTGTTGGAATCTCTACAGT 8618
      |||
Qy      3691 ATTGAGACGAGAACTTAAGAATAGTCTGTTAACTTCTCAATGCCACAGCCATAGCAG 3750
      |||
Db      8619 ATTGAGACGAGAACTTAAGAATAGTCTGTTAACTTCTCAATGCCACAGCCATAGCAG 8678
      |||
Qy      3751 TAGCTGAGGGGACGATAGGGTTATAGAAGTATTACAAGCAGCTATAGAGCTATTCGCC 3810
      |||
Db      8679 TAGCTGAGGGGACGATAGGGTTATAGAAGTATTACAAGCAGCTATAGAGCTATTCGCC 8738
      |||
Qy      3811 ACATACCTAGAAGATAAGACAGAGGCTTGAAAGGATTTTGTATTAAGATGGTGGCAAG 3870
      |||
Db      8739 ACATACCTAGAAGATAAGACAGAGGCTTGAAAGGATTTTGTATTAAGATGGTGGCAAG 8798
      |||
Qy      3871 TGTGTCAAAAAGTAGTGTATTGGATGGCTCTGCTTAAGGAAAGAAAGATGAGACGAGCTGAG 3930
      |||
Db      8799 TGTGTCAAAAAGTAGTGTATTGGATGGCTCTGCTTAAGGAAAGAAAGATGAGACGAGCTGAG 8858
      |||
Qy      3931 CCAGCAGCAGATGGGGTGGGAGCAGTATCTGAGATCTAGCA 3971
      |||
Db      8859 CCAGCAGCAGATGGGGTGGGAGCAGTATCTGAGACCTAGCA 8899
      |||

```

RESULT 6
AA288127
ID AA288127 standard; DNA; 9709 BP.

AC AA288127;
XX 20-APR-2000 (first entry)
DT
XX HIV-1 Nf5/IIIB recombinant clone pNL4-3, complete genome.
DE
XX Lentiviral vector; packaging; gag; pol; gene therapy; infection;
KW gene expression; PCR primer; ss.
XX

```

OS      Human immunodeficiency virus type 1.
OS      Synthetic.
PN      WO200000600-A2.
XX      06-JAN-2000.
PD      26-MAY-1999; 99WO-US11516.
XX      26-MAY-1998; 98US-0086635.
PR      (CHAN/) CHANG L.
XX      Chang L.
PI      WPI: 2000-137067/12.
DR      New packaging vector comprising a nucleotide sequence encoding Gag and
XX      Pol proteins of a reference lentivirus useful for the delivery of
XX      non-lentiviral genes to target cells.
PT      Example: Page 258-263; 311pp; English.
PS      The present invention describes a packaging vector (PV) comprising a
CC      nucleotide sequence encoding Gag and Pol proteins of a reference
CC      lentivirus that differs from the reference lentivirus at least in that:
CC      (a) its major splice donor site is either deleted or is sufficiently
CC      different from the reference lentivirus so that it is not a potential
CC      site for homologous recombination; and (b) it lacks a functional major
CC      packaging signal so that the introduced vector causes the host cell to
CC      produce packaging vector particles comprising functional Gag and Pol
CC      proteins. The vectors are useful for transforming (eukaryotic) cells to
CC      express specific genes at high levels, e.g. for gene therapy. The
CC      improved vectors are safer, yet permit increased efficiency of packaging
CC      the recombinant viral genome and increased long-term gene expression.
CC      These properties are required for gene therapy as a means of treating
CC      infectious and non-infectious diseases. Unlike other retroviruses, the
CC      lentiviruses are able to infect non-dividing cells. The present sequence
CC      represents the HIV-1 Nf5/IIIB recombinant clone pNL4-3, complete genome
CC      which is given in the exemplification of the present invention.
XX      Sequence 9709 BP; 3421 A; 1756 C; 2366 G; 2166 T; 0 other.
SO      Query Match 43.0%; Score 2676.2; DB 21; Length 9709;
      Best Local Similarity 99.9%; Pred. No. 5.2e-314;
      Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy      1291 CCATGAGAGTGAAGAGAGATATCAGCACTTGTGAGATGGGGTGAAGTGGGACCA 1350
      |||
Db      6219 CAATGAGAGTGAAGAGAGATATCAGCACTTGTGAGATGGGGTGAAGTGGGACCA 6278
      |||
Qy      1351 TGTCTCTGGGATATTTGATGATCTGTAGTGTCTACAGAAATTTGTGGTCTACCGTCTATT 1410
      |||
Db      6279 TGTCTCTGGGATATTTGATGATCTGTAGTGTCTACAGAAATTTGTGGTCTACCGTCTATT 6338
      |||
Qy      1411 ATGGGATACCTGTGTGGGAAGAGCAACACACACTCTATTTTGTGATCAGATGCTAAG 1470
      |||
Db      6339 ATGGGATACCTGTGTGGGAAGAGCAACACACACTCTATTTTGTGATCAGATGCTAAG 6398
      |||
Qy      1471 CATATGATACAGAGGTACATATATTTTGGGACACATAGCCGTGTACCCACAGACCCA 1530
      |||
Db      6399 CATATGATACAGAGGTACATATATTTTGGGACACATAGCCGTGTACCCACAGACCCA 6458
      |||
Qy      1531 ACCCACAAGAGTAGTATTTGTAATGTGACAGAAATTTTAACATGTGGAATAATGACA 1590
      |||
Db      6459 ACCCACAAGAGTAGTATTTGTAATGTGACAGAAATTTTAACATGTGGAATAATGACA 6518
      |||
Qy      1591 TGTGTAACACAGATGATGAGATATATACAGTTTATGGGATCAAAAGCCTAAAGCATGTG 1650
      |||
Db      6519 TGTGTAACACAGATGATGAGATATATACAGTTTATGGGATCAAAAGCCTAAAGCATGTG 6578
      |||
Qy      1651 TAAATTAACCCCACTGTGTAGTTTAAAGTGACAGATTTGAAGATGATGATCTATA 1710
      |||

```

Dd	6579	TAATAATTACCCCACTGCTGTGTAGTTTAAAGTCAGCTAGTTTAAAGATATGATACTATATA	6638
Qy	1711	CCAAATAGTAGTACGGGGGGAATGATTAATGAGAAAGAGAGATAAAAAAGCTGCTTTCA	1770
Dd	6639	CCAAATAGTAGTACGGGGGGAATGATTAATGAGAAAGAGAGATAAAAAAGCTGCTTTCA	6698
Qy	1771	ATATCAGACACAGGATATAGATATAGAGGGGAGAAAGAAATATGCACTCTTTATTAACCTG	1830
Dd	6699	ATATCAGACACAGGATATAGAGATATAGAGGGGAGAAAGAAATATGCACTCTTTATTAACCTG	6758
Qy	1831	ATATAGTACCAATAGATTAATACAGCTATAGGTTGATTAAGTTGTAAACCTCAGTCATTA	1890
Dd	6759	ATATAGTACCAATAGATTAATACAGCTATAGGTTGATTAAGTTGTAAACCTCAGTCATTA	6818
Qy	1891	CACAGGCGCTGCCAAAGGATATCCTTTGAGCCAAATTCCTCATTTATTTGCCCCGCTG	1950
Dd	6819	CACAGGCGCTGCCAAAGGATATCCTTTGAGCCAAATTCCTCATTTATTTGCCCCGCTG	6878
Qy	1951	GTTTTGGCATTCTTAAATAGTAAATTAAGAGCTTCAATGGAAGCAAGCAATGTATCAATG	2010
Dd	6879	GTTTTGGCATTCTTAAATAGTAAATTAAGAGCTTCAATGGAAGCAAGCAATGTATCAATG	6938
Qy	2011	TCAGACACGTATCAATGTATACATGGAATCAGGCCAGTAGTATCACTCAACTGCTGTAA	2070
Dd	6939	TCAGACACGTATCAATGTATACATGGAATCAGGCCAGTAGTATCACTCAACTGCTGTAA	6998
Qy	2071	ATGGCAGCTGCAGAAAGATGATGTAGTAATTTATCATCTGCCAATTTCCAGACAACTGA	2130
Dd	6999	ATGGCAGCTGCAGAAAGATGATGTAGTAATTTATCATCTGCCAATTTCCAGACAACTGA	7058
Qy	2131	AAACCATATATAGTACAGCTGACGACATCTGTAGAAATTAATTTGACAAAGCCCAACACA	2190
Dd	7059	AAACCATATATAGTACAGCTGACGACATCTGTAGAAATTAATTTGACAAAGCCCAACACA	7118
Qy	2191	ATACACAGAAAAGATGTCGGTCCAGAGAGGGGACAGGAGAGACTTTGTATACAAATAGAA	2250
Dd	7119	ATACACAGAAAAGATGTCGGTCCAGAGAGGGGACAGGAGAGACTTTGTATACAAATAGAA	7178
Qy	2251	AAATPAGAAATATGAGACACAGACATTTGTAACTTAAGTAGAGCAAAATGGAATGCCACTT	2310
Dd	7179	AAATPAGAAATATGAGACACAGACATTTGTAACTTAAGTAGAGCAAAATGGAATGCCACTT	7238
Qy	2311	TAAACAGATATGCTAGACAAATTAAGAGAACTTGGAAATTAATTAACAAATATCTTTA	2370
Dd	7239	TAAACAGATATGCTAGACAAATTAAGAGAACTTGGAAATTAATTAACAAATATCTTTA	7298
Qy	2371	AGCAATCCTCAGAGAGGGGACCCAGAAATTTGTAAAGCAGAGTTTAAATTTGGAGGGGAAT	2430
Dd	7299	AGCAATCCTCAGAGAGGGGACCCAGAAATTTGTAAAGCAGAGTTTAAATTTGGAGGGGAAT	7358
Qy	2431	TTTTCTACTGTAAATTTCAACACACTGTTTAATAGTACTTGTTTAATAGTACTTGGAGTA	2490
Dd	7359	TTTTCTACTGTAAATTTCAACACACTGTTTAATAGTACTTGTTTAATAGTACTTGGAGTA	7418
Qy	2491	CTGAAGGCTCAAAATTAACACTGAAGAGAGTGACAAATCACACCTCCCATGGAATTAAC	2550
Dd	7419	CTGAAGGCTCAAAATTAACACTGAAGAGAGTGACAAATCACACCTCCCATGGAATTAAC	7478
Qy	2551	AAATTTATTAACATGTGCGAGGAAGTAGAAAAAGCAATGTATGCCCTCCCATCAGTGGAC	2610
Dd	7479	AAATTTATTAACATGTGCGAGGAAGTAGAAAAAGCAATGTATGCCCTCCCATCAGTGGAC	7538
Qy	2611	AAATTTAGATGTTCAATTAATTTCTGGGCTGCTATTAACAAGAGATGGTGTATATPACA	2670
Dd	7539	AAATTTAGATGTTCAATTAATTTCTGGGCTGCTATTAACAAGAGATGGTGTATATPACA	7598
Qy	2671	ACAATGGGCTCGAAGATTTAGACCTGAGAGGCGCATATGAGGGACAATTTGGAAGAGTG	2730
Dd	7599	ACAATGGGCTCGAAGATTTAGACCTGAGAGGCGCATATGAGGGACAATTTGGAAGAGTG	7658
Qy	2731	AATTAATTAATATATAAGTAGTAAAAATTAACCATTAAGAGTAGACCCCAAGAGGCAA	2790
Dd	7659	AATTAATTAATATATAAGTAGTAAAAATTAACCATTAAGAGTAGACCCCAAGAGGCAA	7718

QY	2791	AGAGAAGTGTGTCAGAGAGAAAAAGACACTGCGAATPAGAGCTTTGTTCTTTGGGT	2850
Db	7719	AGAGAAAGTGTGTGCACAGAGAAAAAGACACTGTGGGAATAGAGCTTTGTTCTTTGGGT	7778
QY	2851	TTCTTGGGAGCAGCAGGAAGCACTATGTGGGCTGCACGTCAATGACGGTGAAGAGGCA	2910
Db	7779	TTCTTGGGAGCAGCAGGAAGCACTATGTGGCTGCACGTCAATGACCGTGAAGGTCGA	7838
QY	2911	GACAATTTATTTGTGATATATAGTGCACAGCAGAACAAATTTGCTGAGGGCTATTGAGGGC	2970
Db	7839	GACAATTTATTTGTGATATATAGTGCACAGCAGAACAAATTTGCTGAGGGCTATTGAGGGC	7898
QY	2971	AACGACATCTGTTCACATCTACAGCTGTGGGCAATCAACAGCTCCACGCAACAAATCTGG	3030
Db	7899	AACGACATCTGTTCACATCTACAGCTGTGGGCAATCAACAGCTCCACGCAACAAATCTGG	7958
QY	3031	CTGTGGAAGAATACCTTAAAGATCAACAGCTCCTGTGGGATTTTGGGTTGCTCTGGAAC	3090
Db	7959	CTGTGGAAGAATACCTTAAAGATCAACAGCTCCTGTGGGATTTTGGGTTGCTCTGGAAC	8018
QY	3091	TCATTTGCACCACTGCTGTGCTCTGGAAATGCTAGTTGGAATTAATCTCTGCAACGA	3150
Db	8019	TCATTTGCACCACTGCTGTGCTCTGGAAATGCTAGTTGGAATTAATCTCTGCAACGA	8078
QY	3151	TTTGGAAATTAACATGACCTGGATGGAGTGGGACAGAGAATTAACAAATTAACACACTTAA	3210
Db	8079	TTTGGAAATTAACATGACCTGGATGGAGTGGGACAGAGAATTAACAAATTAACACACTTAA	8138
QY	3211	TACACTCCTTTAATTGAAGATGCGCAAAACCCAGCAAGAAAAAGATGACAGAATTTATGG	3270
Db	8139	TACACTCCTTTAATTGAAGATGCGCAAAACCCAGCAAGAAAAAGATGACAGAATTTATGG	8198
QY	3271	AATAGATTAATATGGGCAAGTTTGTGAAATTTGTTTAACTAACAATTTGGCTGTGATA	3330
Db	8199	AATAGATTAATATGGGCAAGTTTGTGAAATTTGTTTAACTAACAATTTGGCTGTGATA	8258
QY	3331	TAAATATTATTCATPATGATATGATGAGAGGCTTGGTAGTTTAAAGAAATGATTTTGTGTAC	3390
Db	8259	TAAATATTATTCATPATGATATGATGAGAGGCTTGGTAGTTTAAAGAAATGATTTTGTGTAC	8318
QY	3391	TTTCTATAGGAATPAGAGTATGAGGCAAGGAAATTCACCACTTATCGTTGACGCCACCTCC	3450
Db	8319	TTTCTATAGGAATPAGAGTATGAGGCAAGGAAATTCACCACTTATCGTTGACGCCACCTCC	8378
QY	3451	CAATCCGAGAGGAGCCCGACAGGCCCGGAAGAAATAGAAAGAAAGAGTGGAGAGAGACA	3510
Db	8379	CAATCCGAGAGGAGCCCGACAGGCCCGGAAGAAATAGAAAGAAAGAGTGGAGAGAGACA	8438
QY	3511	GAGACAGATCCATTCGATTTAGTATACAGGATCCTTAGACCTTATCTGAGCAGCATCTGCGGA	3570
Db	8439	GAGACAGATCCATTCGATTTAGTATACAGGATCCTTAGACCTTATCTGAGCAGCATCTGCGGA	8498
QY	3571	GCGTGTGCTCTTCAGCTACACACCGGTTAGAGACTTACTCTTGATTTAAGCAGAGATTTG	3630
Db	8499	GCGTGTGCTCTTCAGCTACACACCGGTTAGAGACTTACTCTTGATTTAAGCAGAGATTTG	8558
QY	3631	TGGAATCTTGGGACGCGAGGGGTGGGAAGCCCTCAATATATTTGTGGAATCTCTTACAGT	3690
Db	8619	TGGAATCTTGGGACGCGAGGGGTGGGAAGCCCTCAATATATTTGTGGAATCTCTTACAGT	8678
QY	3691	ATTGGAGTTCAGGAACCTTAAGAAATATGTCGTTAACTGTCTCAATGCTCCACAGCAATPAG	3750
Db	8619	ATTGGAGTTCAGGAACCTTAAGAAATATGTCGTTAACTGTCTCAATGCTCCACAGCAATPAG	8678
QY	3751	TAGCTGAGGGGACAGATAGGGTTATATGAGATATTAACAAGCACTTATAGAGCTATTTGCC	3810
Db	8679	TAGCTGAGGGGACAGATAGGGTTATATGAGATATTAACAAGCACTTATAGAGCTATTTGCC	8738
QY	3811	ACATACCTAGAGAAATPAGACAGGGCTTGGAAAGATTTTGTATPAAAGTGGGTGGCAAG	3870
Db	8739	ACATACCTAGAGAAATPAGACAGGGCTTGGAAAGATTTTGTATPAAAGTGGGTGGCAAG	8798

Qy	Db	Qy	Db
3871	TGGCGAAAAGTACTGATTTGGATTTGGCCCTGCTCTAAGGGGAAATAATGAGACGAGCTGAG	3930	
8799	TGGCGAAAAGTACTGATTTGGATTTGGCCCTGCTCTAAGGGGAAATAATGAGACGAGCTGAG	8858	
3931	CCAGCAGCAGATGGGGGTGGGAGCAGTATCTTCGAGATCTAGA	3971	
8859	CCAGCAGCAGATGGGGGTGGGAGCAGTATCTTCGAGACCTTACA	8899	

RESULT 7	
ABQ76897	
ID	ABQ76897 standard; DNA; 9709 BP.
XX	
AC	ABQ76897;
XX	
DT	11-MAR-2003 (first entry)
XX	
DE	HIV-1 NL4-3 gp120 DNA SEQ ID 1.
XX	
KW	gp120; interaction; co-receptor; CXCR4; CCR5; refractive index; V3 loop; 7-helix transmembrane receptor; glycopeptide; virucide; anti-HIV;
KW	HIV infection; ds.
XX	
OS	Human immunodeficiency virus type 1.

PN	Key	Location/Qualifiers
DE10113042-A1.	CDS	6221..8785
		/*tag= a
		/product= "gp120"

PD	26-SEP-2002.
XX	
PF	09-MAR-2001; 2001DE-1013042
XX	
PR	09-MAR-2001; 2001DE-1013042.

PA (NOCH-) NOCHT INST TROPENMEDIZIN BERNHARD.

PI Schreiber M, Selfert A, meyer B;

DR WPI; 2002-752120/82.
DR P-PSDB; ABG73663.

PT Identifying compounds that modify interaction of gp120 and
PT co-receptors, useful potentially for treating human immune deficiency
PT virus infection, also new peptides -

PS Disclosure; Page 29-33; 68pp; German.

CC This invention describes novel substances that modify the interaction
CC between the gp120 protein of human immunodeficiency virus (HIV), or its
CC fragments, with the co-receptors CXCR4, CCR5 and/or other 7-helix
CC transmembrane receptors for HIV. The method comprises (a) immobilizing a
CC ligand for the co-receptor on a gold surface; (b) contacting the ligand
CC with suspended cells that express the co-receptor; and (c) determining
CC interaction by measuring the refractive index (RI) by plasmon resonance.
CC The procedure is repeated using cells that have been incubated with a
CC test compound, and this is identified if RI is lower for cells
CC preincubated with it. This ligand is a linear or cyclic (glyco)peptide
CC that includes the amino acid sequence of an HIV V3 loop (including
CC flanking Cys). The products of the invention have virucide and anti-HIV
CC (human immunodeficiency virus) activity and are useful for prevention
CC and/or treatment of HIV infection. This sequence encodes the HIV type 1
CC gp120 described in the disclosure of this invention.

SQ Sequence 9709 BP; 3421 A; 1756 C; 2366 G; 2166 T; 0 other;

Query Match	43.08;	Score 2676.2;	DB 24;	Length 9709;
Best Local Similarity	99.98;	Pred. No. 5.2e-314;		
Matches 2678; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

OY	1291	CCATGAGAGTGAAGAGAAAGTATCACACACTTGTGGAGATGGGGGTGGAAATGGGGCACC	1350
Db	6219	CAATGAGAGTGAAGGAAAGTATCACACACTTGTGGAGATGGGGGTGGAAATGGGGCACC	6278
OY	1351	TGCTCCCTTGGGATATTGATGATCTGTAGTGCATACAGAAAAATTGGTGGGTACACCGCTATT	1410
Db	6279	TGCTCCCTTGGGATATTGATGATCTGTAGTGCATACAGAAAAATTGGTGGGTACACAGCTATT	6338
OY	1411	ATGGGGTACTGTGTGGGAAGGAAGCACCACCACTCTATTGTGTGCATCAGATGCTAAAG	1470
Db	6339	ATGGGGTACTGTGTGGGAAGGAAGCACCACCACTCTATTGTGTGCATCAGATGCTAAAG	6398
OY	1471	CATATGATACAGAGTACATATATGTTTGGGCCACACATCGCTGTATCCACAGACCCA	1530
Db	6399	CATATGATACAGAGTACATATATGTTTGGGCCACACATCGCTGTATCCACAGACCCA	6458
OY	1531	ACCCACAAGAAAGTAAATTTGCTAAATGTACACGAAAAATTTTAACTATGTGGAAAAATGACA	1590
Db	6459	ACCCACAAGAAAGTAAATTTGCTAAATGTACACGAAAAATTTTAACTATGTGGAAAAATGACA	6518
OY	1591	TGCTAGAACAGATGCATAGAGATATATAGTTTATGGATCAAAAGCCTTAAAGCCATGTG	1650
Db	6519	TGCTAGAACAGATGCATAGAGATATATAGTTTATGGATCAAAAGCCTTAAAGCCATGTG	6578
OY	1651	TAAATTTAACCCACTCTGTGTATTAGTTTAAAGTCACATGATTTGAAGATGATACTAATA	1710
Db	6579	TAAATTTAACCCACTCTGTGTATTAGTTTAAAGTCACATGATTTGAAGATGATACTAATA	6638
OY	1711	CCATATGTAATACCGGAGATATATATGTGAGAAAGGAGATTAAAAAATCTGCTTTCA	1770
Db	6639	CCATATGTAATACCGGAGATATATATGTGAGAAAGGAGATTAAAAAATCTGCTTTCA	6698
OY	1771	ATATCAGCACAAGCATATAGATATAGGTGAGAAAGAAATATCATTTCTTTTATAACTTG	1830
Db	6699	ATATCAGCACAAGCATATAGATATAGGTGAGAAAGAAATATCATTTCTTTTATAACTTG	6758
OY	1831	ATATATGATACCAATATAGATATATACAGCTATAGTTTGAATGTTTAACACCTCAGTCATTA	1890
Db	6759	ATATATGATACCAATATAGATATATACAGCTATAGTTTGAATGTTTAACACCTCAGTCATTA	6818
OY	1891	CACAGGCTCTCCAAAAGTATCTTTGAGCCCAATTCCTATATATTGTGTGCCGGCTG	1950
Db	6819	CACAGGCTCTCCAAAAGTATCTTTGAGCCCAATTCCTATATATTGTGTGCCGGCTG	6878
OY	1951	GTTTTGGGATTCTTAAATGTAAATTAATTAAGCGTTCAATGGAACAGACCATGTACAATG	2010
Db	6879	GTTTTGGGATTCTTAAATGTAAATTAATTAAGCGTTCAATGGAACAGACCATGTACAATG	6938
OY	2011	TCAGCAGATACATATGATACATGAAATCAGGCCAGCTAGTATCAACTCACTGCTTTAA	2070
Db	6939	TCAGCAGATACATATGATACATGAAATCAGGCCAGCTAGTATCAACTCACTGCTTTAA	6998
OY	2071	ATGGCAGTCTAGCAGAAAGAAAGATGTAGTAAATAGATCTGCCAATTTACAGACAATGCTA	2130
Db	6999	ATGGCAGTCTAGCAGAAAGAAAGATGTAGTAAATAGATCTGCCAATTTACAGACAATGCTA	7058
OY	2131	AAACCAATTAATAGTACAGCTTAACACATCTGTGAATAATTAATTGTCAAGAACCCACAACA	2190
Db	7059	AAACCAATTAATAGTACAGCTTAACACATCTGTGAATAATTAATTGTCAAGAACCCACAACA	7118
OY	2191	ATACAGAAAAAAGTATTCCTATCCAGAGGGGACAGGAGAGCATTTGTTACAAATAGAA	2250
Db	7119	ATACAGAAAAAAGTATTCCTATCCAGAGGGGACAGGAGAGCATTTGTTACAAATAGAA	7178
OY	2251	AAATAGCAATATAGACACAAGCATTTGTACATTTAGTAGACCAAAATGGATGCGACTT	2310
Db	7179	AAATAGCAATATAGACACAAGCATTTGTACATTTAGTAGACCAAAATGGATGCGACTT	7238
OY	2311	TAAACAGATAGCTAGCAATTAAGGAACAATTTGGAAATATATTAACCAATTAATCTTTA	2370
Db	7239	TAAACAGATAGCTAGCAATTAAGGAACAATTTGGAAATATATTAACCAATTAATCTTTA	7298
OY	2371	AGCAATCTCTAGAGAGGGGACCGAATAATTGTAAACGACAGTTTAAATTGTGAGGGGAAT	2430

```

Db      7299 AGCAATCCTCAGAGGGGAGCCGAAATTTGAACGCACAGTTTTTAATTTGGAGGGGAAT 7358
Qy      2431 TTTTCACTGTAATTCACACACACTGTTTAACTAGTACTTGGTTAAATAGTACTGGGTA 2490
Db      7359 TTTTCACTGTAATTCACACACACTGTTTAACTAGTACTTGGTTAAATAGTACTGGGTA 7418
Qy      2491 CTGAGGGTCAAAATACACTGAAGAGTAGACACAAATCACACTCCATGCGAATATAAC 2550
Db      7419 CTGAGGGTCAAAATACACTGAAGAGTAGACACAAATCACACTCCATGCGAATATAAC 7478
Qy      2551 AATTATTAACATGTGGCAGAGTAGAAGAAAGCAATGTATGCCCTCCATCAGTGGAC 2610
Db      7479 AATTATTAACATGTGGCAGAGTAGAAGAAAGCAATGTATGCCCTCCATCAGTGGAC 7538
Qy      2611 AATTAGATGTTCATCAAAATATTACTGCGCTGCTATTACAAGAGATGGTATATACA 2670
Db      7539 AATTAGATGTTCATCAAAATATTACTGCGCTGCTATTACAAGAGATGGTATATACA 7598
Qy      2671 ACAATGGGTCGAGATCTTCAGACCTTGAGAGAGCGATATGAGGGACAAATTTGAGAAAGTG 2730
Db      7599 ACAATGGGTCGAGATCTTCAGACCTTGAGAGAGCGATATGAGGGACAAATTTGAGAAAGTG 7658
Qy      2731 AATTATTAATATTAAGTAACTAAATTTGAACCATTTAGAGATGACACCAAGCAAGCAA 2790
Db      7659 AATTATTAATATTAAGTAACTAAATTTGAACCATTTAGAGATGACACCAAGCAAGCAA 7718
Qy      2791 AGAGAAAGTGTGTCACAGAGAAAGAAAGAGAGAGTGGGAAATAGAGGCTTGTCTCTGGGT 2850
Db      7719 AGAGAAAGTGTGTCACAGAGAAAGAAAGAGAGAGTGGGAAATAGAGGCTTGTCTCTGGGT 7778
Qy      2851 TCTTGGGAGCAGCAGAGAAAGCAATATGGGCTGCACGTCATATGACGCTGACGGTACAGCCA 2910
Db      7779 TCTTGGGAGCAGCAGAGAAAGCAATATGGGCTGCACGTCATATGACGCTGACGGTACAGCCA 7838
Qy      2911 GACATATATGTCTGATATATAGTGTGAGCAGAGAAACAAATTTGCTAGGGCTATTAGAGCCG 2970
Db      7839 GACATATATGTCTGATATATAGTGTGAGCAGAGAAACAAATTTGCTAGGGCTATTAGAGCCG 7898
Qy      2971 AACGACATCTGTTGCACTCACTACAGCTGGGGCATCAAAAGCTCCAGCAAGAAATCCTGG 3030
Db      7899 AACGACATCTGTTGCACTCACTACAGCTGGGGCATCAAAAGCTCCAGCAAGAAATCCTGG 7958
Qy      3031 CTGTGAAAGATACCTTAAAGGATCACAGCTCTGGGGATTTGGGGTGTGCTTGGAAAC 3090
Db      7959 CTGTGAAAGATACCTTAAAGGATCACAGCTCTGGGGATTTGGGGTGTGCTTGGAAAC 8018
Qy      3091 TCATTTGCACCACTGCTGTGCTTGGAAATGCTAGTTGGAGTAAATTAATCTCTGGAAACGA 3150
Db      8019 TCATTTGCACCACTGCTGTGCTTGGAAATGCTAGTTGGAGTAAATTAATCTCTGGAAACGA 8078
Qy      3151 TTTGGAAATACATGACCTGGATGGAGTGGGAGAGAAATTAACAAATTTACCAAGCTTAA 3210
Db      8079 TTTGGAAATACATGACCTGGATGGAGTGGGAGAGAAATTAACAAATTTACCAAGCTTAA 8138
Qy      3211 TACACTCTCTTAATTGAAGAAATCGCAAAACAGCAGAGAAAGAAATGAACAAGAAATTTATGG 3270
Db      8139 TACACTCTCTTAATTGAAGAAATCGCAAAACAGCAGAGAAAGAAATGAACAAGAAATTTATGG 8198
Qy      3271 AATTAGATAAATGGGCAAGTTTGTGAATTGGTTTAAATTAACAAATTTGGCTGTGATATA 3330
Db      8199 AATTAGATAAATGGGCAAGTTTGTGAATTGGTTTAAATTAACAAATTTGGCTGTGATATA 8258
Qy      3331 TAAATATATTCATTAATATAGTAGAGAGCTTGGATTTAAAGAAATTTTCTCTGTAC 3390
Db      8259 TAAATATATTCATTAATATAGTAGAGAGCTTGGATTTAAAGAAATTTTCTCTGTAC 8318
Qy      3391 TTTCTATAGTAGAATAGAGATTAGCAGAGGATATTACCATTAATCTTTACAGACCACTCC 3450
Db      8319 TTTCTATAGTAGAATAGAGATTAGCAGAGGATATTACCATTAATCTTTACAGACCACTCC 8378
Qy      3451 CAATCCGAGGGGAGCCGACAGGCGCGAAGGAATAGAAGAAAGTGTGAGAGAGAGACA 3510

```

```

Db      8379 CAATCCGAGGGGAGCCGACAGGCGCGAAGGAATAGAGAAAGGTGTGAGAGAGACA 8438
Qy      3511 GAGACAGATCCATTGCATTGTAGACGATCCTTAGACACTTATCTGGAGATCTGCGGA 3570
Db      8439 GAGACAGATCCATTGCATTGTAGACGATCCTTAGACACTTATCTGGAGATCTGCGGA 8498
Qy      3571 GCCTGTGCTCTTCACTACCTACACCGCTTGAGAGACTTACTCTGATTGTAAAGAGATTG 3630
Db      8499 GCCTGTGCTCTTCACTACCTACACCGCTTGAGAGACTTACTCTGATTGTAAAGAGATTG 8558
Qy      3631 TGGAACTCTTGGAGCGAGGGGGTGGGAACCCCTCAAAATTTGTGGAAATCTCTACGT 3690
Db      8559 TGGAACTCTTGGAGCGAGGGGGTGGGAACCCCTCAAAATTTGTGGAAATCTCTACGT 8618
Qy      3691 ATTGAGATCAGAACTAAAGAAATAGTCTTAACTTGTCTCAATGCGCACAGCATATGACAG 3750
Db      8619 ATTGAGATCAGAACTAAAGAAATAGTCTTAACTTGTCTCAATGCGCACAGCATATGACAG 8678
Qy      3751 TAGCTGAGGGGACAGATAGGGTTATAGAACTATTACAAAGCAGCTTATAGAGTATTCGCC 3810
Db      8679 TAGCTGAGGGGACAGATAGGGTTATAGAACTATTACAAAGCAGCTTATAGAGTATTCGCC 8738
Qy      3811 ACATFACTAAGAAATTAAGACAGAGGCTTGGAAAGANTTTGCTATAAGATGGGTGGCAAG 3870
Db      8739 ACATFACTAAGAAATTAAGACAGAGGCTTGGAAAGANTTTGCTATAAGATGGGTGGCAAG 8798
Qy      3871 TGGTCAAAAGTATGATGATTGGATGCGCTGTGTAAAGGAAAGATTAAGATGACAGAGCTGAG 3930
Db      8799 TGGTCAAAAGTATGATGATTGGATGCGCTGTGTAAAGGAAAGATTAAGATGACAGAGCTGAG 8858
Qy      3931 CCAGCAGCAGATGGGGTGGGAGCATATCTCGAGATCTAGA 3971
Db      8859 CCAGCAGCAGATGGGGTGGGAGCATATCTCGAGATCTAGA 8899

RESULT 8
ABS57823
ID      ABS57823 standard; DNA: 9709 BP.
XX
AC      ABS57823;
XX
DT      03-FEB-2003 (first entry)
XX
DE      HIV-1 genome clone pNL4-3.
XX
KW      HIV-1; ss: genome; retroviral vector; gag; pol; envelope protein;
KW      replication-competent retroviral vector; RIR; pcMV-deltaE-P;
KW      pcMV-deltaE-P; pcDNA3-HIVgag-RRE; pcMV-deltaE-P;
XX
OS      Human immunodeficiency virus type 1.
XX
PN      US6451304-B1.
XX
PD      17-SEP-2002.
XX
PE      09-MAR-1999; 99US-0265013.
XX
PR      09-MAR-1998; 98US-077625P.
XX
PA      (REGC ) UNIV CALIFORNIA.
PI      Friedmann T, Miyahara A;
XX
DR      WPI: 2003-074128/07.
XX
PT      Producing replication-competent retrovirus vectors, by transfecting
PT      cells with a provirus plasmid that encodes gag, a provirus plasmid that
PS      encodes pol, and an envelope protein encoding construct -
XX
XX      Example 2; Page -: 16pp; English.
XX
CC      The invention relates to producing replication-competent retrovirus
CC      vectors (RIR), involving transfecting cells with a first provirus plasmid

```


QY	3151	TTTGGAAATAAACATGACCTGGATGGATGGGACAGAGAAATTAACAATTACCAAGCTTAA	33210	FT	Key	Location/Qualifiers
Db	8079	TTTTGGAAATAAACATGACCTGGATGGATGGGACAGAGAAATTAACAATTACCAAGCTTAA	81338	FT	repeat_region	1..634
QY	3211	TACACTCCTTAATTGGAAGATCGCAAAACACAGCAAGAAAAGATGAACAAATTTATGG	3270	FT		/*tag= a
Db	8139	TACACTCCTTAATTGGAAGATCGCAAAACACAGCAAGAAAAGATGAACAAATTTATGG	8198	FT	repeat_unit	/rpl_type= TERMINAL
QY	3271	AATTAGATAAATGGGCACTTTGTGGAAATGGTTTAAACATTAACAAATTTGGCTGTGATA	3330	FT		/note= "5'LTR"
Db	8199	AATTAGATAAATGGGCACTTTGTGGAAATGGTTTAAACATTAACAAATTTGGCTGTGATA	8258	FT		456..548
QY	3331	TAAATTTATCATATGATAGTATAGAGGCTGTAGTGTAGATATAGTTTTCCTGTAC	3390	FT	GC_signal	/*tag= b
Db	8259	TAAATTTATCATATGATAGTATAGAGGCTGTAGTGTAGATATAGTTTTCCTGTAC	8318	FT		/standard_name= R
QY	3391	TTTTCATATGTAATAGACTTAGCGCAGGAGATATTCACATTAATGCTTTCAAGCCACCTCC	3450	FT	GC_signal	/*tag= c
Db	8319	TTTTCATATGTAATAGACTTAGCGCAGGAGATATTCACATTAATGCTTTCAAGCCACCTCC	8378	FT		/standard_name= Spl_binding_site
QY	3451	CAATCCGAGGGGACCCGACAGGCCCGAGAGATAGAAAGAGTGAGAGAGACACA	3510	FT	primer_bind	389..395
Db	8379	CAATCCGAGGGGACCCGACAGGCCCGAGAGATAGAAAGAGTGAGAGAGACACA	8438	FT		/*tag= d
QY	3511	GAGACAGATCCATTGCATTAGTGAACGAGATCCTTAGCACTTATCTGGAGCATCTGGCA	3570	FT		/standard_name= Spl_binding_site
Db	8439	GAGACAGATCCATTGCATTAGTGAACGAGATCCTTAGCACTTATCTGGAGCATCTGGCA	8498	FT		399..407
QY	3571	GCGTGTGCTCTTCAGCTACCAACGCGTTGAGAGACTTACTTGTAAACGAGATG	3630	FT	CDS	/*tag= e
Db	8499	GCGTGTGCTCTTCAGCTACCAACGCGTTGAGAGACTTACTTGTAAACGAGATG	8558	FT		/standard_name= Spl_binding_site
QY	3631	TGGAACCTCTGGAGCGCAGGGGCTGGGAAGCCCTCAATATTTGGTGAATCCTCTACAGT	3690	FT		636..656
Db	8559	TGGAACCTCTGGAGCGCAGGGGCTGGGAAGCCCTCAATATTTGGTGAATCCTCTACAGT	8618	FT	CDS	/*tag= f
QY	3691	ATTGGAGTCAGGAACCTAAAGATATGCTGTTAACTTGCTCATATGCCACAGCCATAGCAG	3750	FT		/standard_name= Lys_tRNA_pbs
Db	8619	ATTGGAGTCAGGAACCTAAAGATATGCTGTTAACTTGCTCATATGCCACAGCCATAGCAG	8678	FT		790..2292
QY	3751	TAGCTGAGGGGACAGATAGGGTTTATAGAAGTATACAGAGCTTATAGACTATTCGCC	3810	FT	exon	/*tag= g
Db	8679	TAGCTGAGGGGACAGATAGGGTTTATAGAAGTATACAGAGCTTATAGACTATTCGCC	8738	FT		/product= gag
QY	3811	ACATACCTAGAGAATAAGACAGGCGTTGGAAGAGATTTGCTATAAGATGGGTGCACG	3870	FT		/*tag= h
Db	8739	ACATACCTAGAGAATAAGACAGGCGTTGGAAGAGATTTGCTATAAGATGGGTGCACG	8798	FT	exon	/note= "NH2-terminal uncertain"
QY	3871	TGGTCAAAAGATAGTGGATGGTGGCGCTGCTGCTTAAGAGGAAAGATGACAGACGCTGAG	3930	FT		5041..5619
Db	8799	TGGTCAAAAGATAGTGGATGGTGGCGCTGCTGCTTAAGAGGAAAGATGACAGACGCTGAG	8858	FT		/*tag= i
QY	3931	CCAGCAGCAGATGGGCTGGGAGCAGTATCTCGAGATCTAGA	3971	FT	CDS	/product= vif
Db	8859	CCAGCAGCAGATGGGCTGGGAGCAGTATCTCGAGATCTAGA	8899	FT		5559..5849
RESULT 9	AAQ22488	standard; DNA; 9709 BP.		FT		/*tag= j
ID	AAQ22488	standard; DNA; 9709 BP.		FT		/product= vpr
AC	AAQ22488;			FT	exon	/*tag= k
XX	06-JUL-1992	(first entry)		FT		/product= vpu
XX	HIV-1 proviral clone pNL4-3.			FT		8380..6044
XX	AIDS; Acquired Immune Deficiency Syndrome; polymerase chain reaction; PCR; site-directed mutagenesis; retrovirus; null-mutation; human; ss.			FT		/*tag= l
XX	Human immunodeficiency virus.			FT		/note= "full-length tat obtained by splicing"
XX				FT		5969..6044
XX				FT		/*tag= m
XX				FT		/product= rev
XX				FT	exon	/note= "Full-length rev obtained by splicing"
XX				FT		8369..8414
XX				FT		/*tag= n
XX				FT		/product= tat
XX				FT		/note= "see above"
XX				FT	exon	8369..8643
XX				FT		/*tag= O
XX				FT		/product= rev
XX				FT		/note= "see above"
XX				FT	CDS	6221..8785
XX				FT		/*tag= p
XX				FT		/product= env
XX				FT	CDS	8787..9407
XX				FT		/*tag= q
XX				FT		/product= nef
XX				FT		9076..9709
XX				FT	repeat_region	/*tag= r
XX				FT		/rpl_type= TERMINAL
XX				FT		/note= "3'LTR"
XX				FT	repeat_unit	9531..9624
XX				FT		/*tag= s
XX				FT		/standard_name= R
XX				FT		9602..9607
XX				FT		/*tag= t
XX				FT		
XX				FT		
XX				FT		
XX				FT		

Db	6644	TTGGTAGAACAGATCCATGAGGATATTAACGTTTATGTGGATCCAAAGCCCTAAAGCCATGTG	6703
QY	1651	TAAATATTAACCCCACTCTGTGTGTAGTTTAAAGTGCACGTATTTGAAGATGTACTATATA	1710
Db	6704	TAAATATTAACCCCACTCTGTGTGTAGTTTAAAGTGCACGTATTTGAAGATGTACTATATA	6763
QY	1711	CCAATAGTAGTAGCGGGAGAAATGTAAATGAGAAAGAGAGATAAAAAACTGCTCTTCA	1770
Db	6764	CCAATAGTAGTAGCGGGAGAAATGTAAATGAGAAAGAGAGATAAAAAACTGCTCTTCA	6823
QY	1771	ATATCAGCACAGCATTAAGAGATTAAGTGCAGAAAGAAATATGCATTTCTTTATATAACTTG	1830
Db	6824	ATATCAGCACAGCATTAAGAGATTAAGTGCAGAAAGAAATATGCATTTCTTTATATAACTTG	6883
QY	1831	ATATAGTATACCAATATGATTAATACGAGTATAGCTGTATAGTTTATACACCTCAGTCAATTA	1890
Db	6884	ATATAGTATACCAATATGATTAATACGAGTATAGCTGTATAGTTTATACACCTCAGTCAATTA	6943
QY	1891	CACAGGCCCTGCCAAAGGATATCTTTGAGCCCAATTTCCCATACATTAATTTGTGCCCCGGCTG	1950
Db	6944	CACAGGCCCTGCCAAAGGATATCTTTGAGCCCAATTTCCCATACATTAATTTGTGCCCCGGCTG	7003
QY	1951	GTTTTGGATTCCTAAATATGTAATTAATAGACTGCATAGACAGGACCCATGTCCAAATG	2010
Db	7004	GTTTTGGATTCCTAAATATGTAATTAATAGACTGCATAGACAGGACCCATGTCCAAATG	7063
QY	2011	TCACACACAGTATCAATGTATACACATGTGAACAGCCAGTATCAACATCAATGCTGTATTA	2070
Db	7064	TCACACACAGTATCAATGTATACACATGTGAACAGCCAGTATCAACATCAATGCTGTATTA	7123
QY	2071	ATGGCAGTCTTAGCAGAAAGAGATGTAGTAATTTAGATCTGCCAATTTCCACAGACAATGCTA	2130
Db	7124	ATGGCAGTCTTAGCAGAAAGAGATGTAGTAATTTAGATCTGCCAATTTCCACAGACAATGCTA	7183
QY	2131	AAACCATTAATGTATACAGTGAACCATCTGTAGAAATTAATTTGTAACAAGACCACCAACA	2190
Db	7184	AAACCATTAATGTATACAGTGAACCATCTGTAGAAATTAATTTGTAACAAGACCACCAACA	7243
QY	2191	ATTACAGAAAAAGATATCCGTATCCAGAGGGGACAGGAGAGCATTTGTTCATATAGGAA	2250
Db	7244	ATTACAGAAAAAGATATCCGTATCCAGAGGGGACAGGAGAGCATTTGTTCATATAGGAA	7303
QY	2251	AAATAGGAAATATGAGAACACACATTTGTAACATTTAGTAGAGCAAAATGGAATGCCACTT	2310
Db	7304	AAATAGGAAATATGAGAACACACATTTGTAACATTTAGTAGAGCAAAATGGAATGCCACTT	7366
QY	2311	TAAAACAGATAGCTAGCAAAATTTAAGAGAACAAATTTGCAATTAATTAATAAATTAATCTTTA	2370
Db	7364	TAAAACAGATAGCTAGCAAAATTTAAGAGAACAAATTTGCAATTAATTAATAAATTAATCTTTA	7423
QY	2371	AGCAATCTCTCGAGAGGGGACCCCAAAATTTGTAACGACAGTTTAAATTTGTGAGGGGAT	2430
Db	7424	AGCAATCTCTCGAGAGGGGACCCCAAAATTTGTAACGACAGTTTAAATTTGTGAGGGGAT	7483
QY	2431	TTTTCTACTGTAAATTTCAACACACCTGTTTAAATAGTACTTGTTTAAATAGTACTTGGAATA	2490
Db	7484	TTTTCTACTGTAAATTTCAACACACCTGTTTAAATAGTACTTGTTTAAATAGTACTTGGAATA	7543
QY	2491	CTGAAGGGTCAATTAACACTGAAGAACTGACACAAATCACACTCCCATGAGAAATAAAAC	2550
Db	7544	CTGAAGGGTCAATTAACACTGAAGAACTGACACAAATCACACTCCCATGAGAAATAAAAC	7603
QY	2551	AATTTATTAACATGTGCGAGGAAGTAGGAAAAAGCAATGTAATGCCCTCCCATGAGTGAAC	2610
Db	7604	AATTTATTAACATGTGCGAGGAAGTAGGAAAAAGCAATGTAATGCCCTCCCATGAGTGAAC	7663
QY	2611	AAATTTAGATGTTCAATCAATATTAATCTGGGCTGCTATTAACAAGAGATGGTGTATATACA	2670
Db	7664	AAATTTAGATGTTCAATCAATATTAATCTGGGCTGCTATTAACAAGAGATGGTGTATATACA	7722
QY	2671	ACAATGGGTCGAGATCTTCAGACCTGAGGAGCGATATAGGAGCAATTTGGAGAACTG	2730

Db	7724	ACAATGGGTCGGAGATCTTTAGACCTCTGGAGGAGGCCATATGAGGGCAATTGGAGACGTG	7783
Qy	2731	AATTATATATAATAAAGTAGTAAAAATTGAAACCATTTAGAGTAGGACCCACCAGGCAA	2790
Db	7784	AATTATATATAAATAGTAGTAAAAATTGAACCATTTAGAGTAGGACCCACCAGGCAA	7843
Qy	2791	AGAGAAAGTGTGTCCAGAGAGAAAAAGACAGTGGGAATAGGAGCTTTGTCTTGGGT	2850
Db	7844	AGAGAAAGTGTGTCCAGAGAGAAAAAGACAGTGGGAATAGGAGCTTTGTCTTGGGT	7903
Qy	2851	TCTTGGGGGACGACGAGAAACATATGGGGTGCACGTCATATGACGCTGACGGTATACAGGCCA	2910
Db	7904	TCTTGGGGGACGACGAGAAACATATGGGGTGCACGTCATATGACGCTGACGGTATACAGGCCA	7963
Qy	2911	GACATTTATTTGTCTGATATATAGTGCAGCAGCAGAAACAATTTGCTGAGGGCTATTGAGGGCC	2970
Db	7964	GACATTTATTTGTCTGATATATAGTGCAGCAGCAGAAACAATTTGCTGAGGGCTATTGAGGGCC	8023
Qy	2971	AACAGCATCTGTTCGAACTCACAGTCTGGGGCATCAACAGCTCCAGGCAGAAATCTCTGG	3030
Db	8024	AACAGCATCTGTTCGAACTCACAGTCTGGGGCATCAACAGCTCCAGGCAGAAATCTCTGG	8083
Qy	3031	CTGGAGAAAGTACCTTAAGAGATCAACAGCTCCGGGGATTTTGGGGTCTCTCGAANAAC	3090
Db	8084	CTGGAGAAAGTACCTTAAGAGATCAACAGCTCCGGGGATTTTGGGGTCTCTCGAANAAC	8143
Qy	3091	TCATTTTGACCACTGCTGTGCTCTTGGAATGCTAGTTGGAGTAATTAATCTCTGGAACAGA	3150
Db	8144	TCATTTTGACCACTGCTGTGCTCTTGGAATGCTAGTTGGAGTAATTAATCTCTGGAACAGA	8203
Qy	3151	TTTGGAAATACATACCTCGATGGATGGGACAGAGAAATTAACAATTTACACAAGCTTAA	3210
Db	8204	TTTGGAAATACATACCTCGATGGATGGGACAGAGAAATTAACAATTTACACAAGCTTAA	8263
Qy	3211	TACACTCCTTAATTGAAGAAATCGCAAAACAGCAAGAAAGATGAACAAGATTTATTTGG	3270
Db	8264	TACACTCCTTAATTGAAGAAATCGCAAAACAGCAAGAAAGATGAACAAGATTTATTTGG	8323
Qy	3271	AATTAGATAAATGGGGCAAGTTTGTGGAATTTGTTTAAACATTAACAATTTGGCTGTGTATA	3330
Db	8324	AATTAGATAAATGGGGCAAGTTTGTGGAATTTGTTTAAACATTAACAATTTGGCTGTGTATA	8383
Qy	3331	TAAATTTATTCATTAATGATAGTAGGAGGCTTTGGTAGTTTAAGAAATAGTTTTGCTGTAC	3390
Db	8384	TAAATTTATTCATTAATGATAGTAGGAGGCTTTGGTAGTTTAAGAAATAGTTTTGCTGTAC	8443
Qy	3391	TTTCTATATGTCAATATAGATAGGACGAGGATATTCACCATATATGCTTTCAGACCCACTCC	3450
Db	8444	TTTCTATATGTCAATATAGATAGGACGAGGATATTCACCATATATGCTTTCAGACCCACTCC	8503
Qy	3451	CAATCCCGAGGGAGCCGACAGGCGCCGAGAAATAGAAAGAGGAGGAGAGAGACA	3510
Db	8504	CAATCCCGAGGGAGCCGACAGGCGCCGAGAAATAGAAAGAGGAGGAGAGAGACA	8563
Qy	3511	GAGACAGATCCATTGATTTAGTTGAACGGAATCTTAGCACCTTATCTGGAGCATCTGGGA	3570
Db	8564	GAGACAGATCCATTGATTTAGTTGAACGGAATCTTAGCACCTTATCTGGAGCATCTGGGA	8623
Qy	3571	GCTGTGCTCTTACAGCTACACCGCTTGAGAGACTTACTCTTGATTTAGCAGAGATTG	3630
Db	8624	GCTGTGCTCTTACAGCTACACCGCTTGAGAGACTTACTCTTGATTTAGCAGAGATTG	8683
Qy	3631	TGGAACTCTTGGAGCGAGGGGGTGGGAAGCCCTCAAAATTTTGGTGAATCTCTACAGT	3690
Db	8684	TGGAACTCTTGGAGCGAGGGGGTGGGAAGCCCTCAAAATTTTGGTGAATCTCTACAGT	8743
Qy	3691	ATTGAGATCAGAACTAAAGAAATAGTGTCTGTTAACTTGTCAATGCCCACAGCCATAGCAG	3750
Db	8744	ATTGAGATCAGAGACTAAAGAAATAGTGTCTGTTAGCTTGTCAATGCCCACAGCCATAGCAG	8803
Qy	3751	TAGCTGAGGGACAGATATGGGTTATATAAGTATTTACACAGAGCTTATAGAGCTATTTGCC	3810
Db	8804	TAGCTGAGGGACAGATATGGGTTATATAAGTATTTACACAGAGAGGCTTATAGAGCTATTTGCC	8863

QY 2371 AGCAATCCCTCAGAGGGAGCCAGAAATTTGAACGACAGTTTAAATTTGGAGGGGAAAT 2430
 |||||
 Db 7299 AGCAATCCTCAGAGGGAGCCAGAAATTTGAACGACAGTTTAAATTTGGAGGGGAAAT 7358
 |||||
 QY 2431 TTTTCTACTGTAAATTCACACAACTGTTTAATAGTACTGTTTAATAGTACTGAGTA 2490
 |||||
 Db 7359 TTTTCTACTGTAAATTCACACAACTGTTTAATAGTACTGTTTAATAGTACTGAGTA 7418
 |||||
 QY 2491 CTGAAGGTCMAAATTAACACTGAGGAAGTGAACACATACACTCCCATGCGAATTAAC 2550
 |||||
 Db 7419 CTGAAGGTCMAAATTAACACTGAGGAAGTGAACACATACACTCCCATGCGAATTAAC 7478
 |||||
 QY 2551 AATTTATTAACATGTCGAGAGAGTAAGAAAAAGCAATGTATGCCCTCCCATCAATGGAC 2610
 |||||
 Db 7479 AATTTATTAACATGTCGAGAGAGTAAGAAAAAGCAATGTATGCCCTCCCATCAATGGAC 7538
 |||||
 QY 2611 AATTTAGATGTTTCATCAAAATTTACTGGGCTGCTATTAAACAAGAGATGGTGAATACA 2670
 |||||
 Db 7539 AATTTAGATGTTTCATCAAAATTTACTGGGCTGCTATTAAACAAGAGATGGTGAATACA 7598
 |||||
 QY 2671 ACAATGGTCCGAGATCTTCAGACTGAGAGAGGAGATATGAGGCAATTTGGAAGATG 2730
 |||||
 Db 7599 ACAATGGTCCGAGATCTTCAGACTGAGAGAGGAGATATGAGGCAATTTGGAAGATG 7658
 |||||
 QY 2731 AATTTATTAATTAATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 2790
 |||||
 Db 7659 AATTTATTAATTAATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 7718
 |||||
 QY 2791 AGAGAAGATGGTGCAGAGAGAAAAAGAGCAAGTGGGAATAGAGACTTGTTCCTTGGGT 2850
 |||||
 Db 7719 AGAGAAGATGGTGCAGAGAGAAAAAGAGCAAGTGGGAATAGAGACTTGTTCCTTGGGT 7778
 |||||
 QY 2851 TCTTGGGAGCAGCAGGAAGACACTATGGGCTGCACGTCATGACGCTGACGCTACAGGCCA 2910
 |||||
 Db 7779 TCTTGGGAGCAGCAGGAAGACACTATGGGCTGCACGCTCATGACGCTGACGCTACAGGCCA 7838
 |||||
 QY 2911 GACATTTATTTGCTGATATAGTATGACAGCAGCAAAATTTGCTGAGGGCTATTGAGGGCG 2970
 |||||
 Db 7839 GACATTTATTTGCTGATATAGTATGACAGCAGCAAAATTTGCTGAGGGCTATTGAGGGCG 7898
 |||||
 QY 2971 AACAGCATCTGTGGCACTCAACAGTCTGGGGCATCAACAGCTCCAGGCAAGAAATCTGG 3030
 |||||
 Db 7899 AACAGCATCTGTGGCACTCAACAGTCTGGGGCATCAACAGCTCCAGGCAAGAAATCTGG 7958
 |||||
 QY 3031 CTGTGGAAGATACCTAAAGGATCAACAGCTCTGTGGGATTTGGGCTTCTGGAANAAC 3090
 |||||
 Db 7959 CTGTGGAAGATACCTAAAGGATCAACAGCTCTGTGGGATTTGGGCTTCTGGAANAAC 8018
 |||||
 QY 3091 TCATTTGCACCACTGCTGCTGGCTGGAGTGTAGTGAATTAATCTGCGAAGACA 3150
 |||||
 Db 8019 TCATTTGCACCACTGCTGCTGGCTGGAGTGTAGTGAATTAATCTGCGAAGACA 8078
 |||||
 QY 3151 TTTGGAAATTAACATGACCTGATGAGTGGAGAGAGAAATTAACAAATTAACAAGCTTAA 3210
 |||||
 Db 8079 TTTGGAAATTAACATGACCTGATGAGTGGAGAGAGAAATTAACAATTAACAAGCTTAA 8138
 |||||
 QY 3211 TACACTCCTTAATTAAGAAATCGCAAAACGAGCAAGAAAGAAATGACAGAATTTATGG 3270
 |||||
 Db 8139 TACACTCCTTAATTAAGAAATCGCAAAACGAGCAAGAAAGAAATGACAGAATTTATGG 8198
 |||||
 QY 3271 AATTTAGATTAATGGCAAGTTTGGAGTGTGTTTAACATTAACAATTTGGCTGTAATA 3330
 |||||
 Db 8199 AATTTAGATTAATGGCAAGTTTGGAGTGTGTTTAACATTAACAATTTGGCTGTAATA 8258
 |||||
 QY 3331 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3390
 |||||
 Db 8259 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8318
 |||||
 QY 3391 TTTCTATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3450
 |||||
 Db 8319 TTTCTATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8378
 |||||

QY 3451 CAATCCGAGGGAGCCGACAGGCCCCGAGAGATAGAAAGAAAGTGGAGAGAGACA 3510
 |||||
 Db 8379 CAATCCGAGGGAGCCGACAGGCCCCGAGAGATAGAAAGAAAGTGGAGAGAGACA 8438
 |||||
 QY 3511 GAGACAGATTCATTCGATTAATAGTAAGAGATCCTTAAGCACTTATCTGGAGCATTCGCGA 3570
 |||||
 Db 8439 GAGACAGATTCATTCGATTAATAGTAAGAGATCCTTAAGCACTTATCTGGAGCATTCGCGA 8498
 |||||
 QY 3571 GCCCTGGCTCTTCACGCTACCAACCCCTTGAGAGACTTACTCTTGATTGTAACGAGATTTG 3630
 |||||
 Db 8499 GCCCTGGCTCTTCACGCTACCAACCCCTTGAGAGACTTACTCTTGATTGTAACGAGATTTG 8558
 |||||
 QY 3631 TGGAACTCTTGGAGAGCGAGGGGTGGGAAGCCCTCAATATTGGGAATCTCCACAGT 3690
 |||||
 Db 8559 TGGAACTCTTGGAGAGCGAGGGGTGGGAAGCCCTCAATATTGGGAATCTCCACAGT 8618
 |||||
 QY 3691 ATTGGAGTCAGAGACTTAAGAAATAGTGTCTTTAACTTCTCAATGCCACAGCCATAGCAG 3750
 |||||
 Db 8619 ATTGGAGTCAGAGACTTAAGAAATAGTGTCTTTAACTTCTCAATGCCACAGCCATAGCAG 8678
 |||||
 QY 3751 TAGCTGAGGGGACATATAGGCTTATAGAAATTTACAAGCAGCTTATAGACTATTCGCC 3810
 |||||
 Db 8679 TAGCTGAGGGGACATATAGGCTTATAGAAATTTACAAGCAGCTTATAGACTATTCGCC 8738
 |||||
 QY 3811 ACATACCTAGAAAGATTAAGACAGAGGCTTGGAAAGATTTGCTATTAAGATGGGTGGCAAG 3870
 |||||
 Db 8739 ACATACCTAGAAAGATTAAGACAGAGGCTTGGAAAGATTTGCTATTAAGATGGGTGGCAAG 8798
 |||||
 QY 3871 TGGTCAAAAAGTAGTGTATTTGATGGGCTGCTGTAAAGGAAAGAAATGAGAGAGAGCTGAG 3930
 |||||
 Db 8799 TGGTCAAAAAGTAGTGTATTTGATGGGCTGCTGTAAAGGAAAGAAATGAGAGAGAGCTGAG 8858
 |||||
 QY 3931 CCAGCAGCAGATGGGGTGGAGCAGTA 3957
 |||||
 Db 8859 CAAGAAATGGCTAGCAAAAGAGAAAGAA 8885
 |||||

RESULT 12
 AAV74271
 ID AAV74271 standard; DNA; 9213 BP.
 XX
 AC AAV74271;
 XX
 DT 14-JUN-1999 (first entry)
 XX
 DE HTLV-III genomic DNA.
 XX
 KW AIDS; acquired immunodeficiency syndrome; viral infection; env protein;
 KW fusion polypeptide; E' protein; gag protein; pol protein; p' protein;
 KW gp signal peptide; detection; vaccination; etiological agent;
 KW Infection inhibitor; AIDS associated retrovirus; p24; gp41; ss.
 XX
 OS Human lymphotropic virus type III.
 XX
 FH Key Location/Qualifiers
 FT CDS 336..1772
 FT /*tag= a
 FT /product= "gag protein"
 FT 732..1772
 FT /*tag= b
 FT /product= "p24 gag protein"
 FT 1639..4677
 FT /*tag= c
 FT /product= "pol protein"
 FT 4622..5200
 FT /*tag= d
 FT /product= "p' protein"
 FT 5803..8373
 FT /*tag= e
 FT /product= "env protein"
 FT 7336..8373
 FT /*tag= f
 FT /product= "gp41 env protein"

OY		3805	TTCCGCCACATACCCTAGAAAGATTAAACAGCGGCTTGGAAAGCATTTTCCTTAAGATGGGT	3864
Dd		8321	TTCCGCCACATACCTTAGAAGAATTAAGACAGGGCCTTGAAAAGGATTTTGCCTTAAGATGGGT	8380
OY		3865	GCGCAAGTGTCTAAAAAGACTGTGATTGGATGGCGCTGCTCTTAAGGGAAGCAATGACAGCA	3924
Dd		8381	GCGCAAGTGTCTAAAAAGACTGTGATTGGATGGCGCTGCTCTTAAGGGAAGCAATGACAGCA	8440
OY		3925	GCTGAGCCAGCAGCAGATGGGGTGGGAGCACGATATCTCGAATCTAGA	3971
Dd		8441	GCTGAGCCAGCAGCAGATGGGGTGGGAGCACGATCTCTGAGACCTAGA	8487
RESULT 13				
ID	AAV81866		standard; DNA; 9213 BP.	
XX	AAV81866;			
AC	AAV81866;			
XX				
DT	01-JUN-1999	(first entry)		
DE			Nucleic acid sequence of an HIV-1 isolate.	
XX				
KW			Antigenic composition; primate; lentivirus; nef gene; vaccine; infection; AIDS; HIV-1, ss.	
OS			Human immunodeficiency virus type 1.	
PN	US5851813-A.			
PD				
XX	22-DEC-1998.			
PF	27-JAN-1994;	94US-0188583.		
PR	27-JAN-1994;	94US-0188583.		
PR	12-JUL-1990;	90US-0551945.		
XX	09-JUL-1991;	91US-0727494.		
PA	(HARD) HARVARD COLLEGE.			
PI	Destroiers RC;			
DR	WIPI: 1999-080408/07.			
XX	P-FSDB: AAM89322, AAM89323, AAM89324, AAM89325, AAM89326.			
PT	Lentivirus antigenic compositions - containing lentivirus with nef gene deletion			
PS	Disclosure; Flg 2A-R; 93pp; English.			
XX	The invention relates to an antigenic composition comprising an isolated primate lentivirus whose genome contains an engineered non-revertible null mutation in the nef gene, or an infectious DNA clone in a carrier. The antigenic composition is used in vaccines against infection by the lentivirus, e.g. AIDS. The present sequence represents the nucleic acid sequence of an HIV-1 isolate.			
SQ	Sequence 9213 BP; 3297 A; 1655 C; 2218 G; 2043 T; 0 other;			
Query Match				
Best Local Similarity 41.7%; Score 2599.4; DB 20; Length 9213;				
Matches 2640; Conservative 98.3%; Pred. No. 1e-304;				
OY 1291 CCATGAGAGTGAGAGAGATATCAGCACCTTGTGGAGATGGGGGTGCATAATGGGGCACCA 1350				
Dd 5801 CAATGAGAGTGAGAGAGAGAAATATCAGCACCTTGTGAGATGGGGGTGAGATGGGGCACCA 5860				
OY 1351 TGCTCCCTTGGGATATATGATGATCTCTGATGCTACAGAAAATTCTGGGTACCGCTTATT 1410				
Dd 5861 TGCTCCCTTGGGATATGATGATCTCTGATGCTACAGAAAATTCTGGGTACAGCTTATT 5920				
OY	1411 ATGGGGTACCTGTGTGNAAGNACCAACCACTATTTTGTGCAATCAGATGCTAAAG	1470		

|||||
Db 5921 ATGGGTACCTGTGTGGGAAGCAACACCACTATTTTGTGATCGATGCTAAAG 5980
QY 1471 CATTATACAGAGGTGACATTAATGTTTGGGCCACATGCTGTGTAACCCACAGACCCA 1530
Db 5981 CATATATACAGAGGTGACATTAATGTTTGGGCCACATGCTGTGTAACCCACAGACCCA 6040
QY 1531 ACCCACAAGAGTAGTATTGGTAAATGTGCACAGAAAATTTTAACATGTGGAAAATGACA 1590
Db 6041 ACCCACAAGAGTAGTATTGGTAAATGTGCACAGAAAATTTTAACATGTGGAAAATGACA 6100
QY 1591 TGGTAGACAGATGATGAGGATTAATCATGTTTATGGGATCAAAGCCTTAAGCCATGTG 1650
Db 6101 TGGTAGACAGATGATGAGGATTAATCATGTTTATGGGATCAAAGCCTTAAGCCATGTG 6160
QY 1651 TAAATTAACCCACCTGTGTGTTAAAGTGCACTGATTTGAAGAAATGATCTATA 1710
Db 6161 TAAATTAACCCACCTGTGTGTTAAAGTGCACTGATTTGAAGAAATGATCTATA 6220
QY 1711 CCAATAGTAGTACGGGAGAAATGATATATGAGAAAGAGAGATAAAAACTGCTTTCA 1770
Db 6221 CCAATAGTAGTACGGGAGAAATGATATATGAGAAAGAGAGATAAAAACTGCTTTCA 6280
QY 1771 ATATCAGCAGACATTAAGAGATTAAGGTGAGAAAGAAATATGCAATCTTTTATTAACCTTG 1830
Db 6281 ATATCAGCAGACATTAAGAGATTAAGGTGAGAAAGAAATATGCAATCTTTTATTAACCTTG 6340
QY 1831 ATATAGTACCAATAGATATA-----TACCACTATAGTGTGATAGTGTAAACCCAG 1884
Db 6341 ATATTAATCCAAATAGATATAGATCTACACGTATACCTTGTGACAAATGTAACACCTCAG 6400
QY 1885 TCATTAACACAGGCTGTCCAAAGGTATCCTTTGAGCCAAATCCCATCATTAATTTGCCCC 1944
Db 6401 TCATTAACACAGGCTGTCCAAAGGTATCCTTTGAGCCAAATCCCATCATTAATTTGCCCC 6460
QY 1945 CGGCTGGTTTGGGATTTCTAAAAATGTAATTAAGACCTTCAATGGAACGAGACATGTA 2004
Db 6461 CGGCTGGTTTGGGATTTCTAAAAATGTAATTAAGACCTTCAATGGAACGAGACATGTA 6520
QY 2005 CAAATGTCAGCAGTACATGTAACATGTAACATGTAACATGTAACATGTAACATGTAAC 2064
Db 6521 CAAATGTCAGCAGTACATGTAACATGTAACATGTAACATGTAACATGTAACATGTAAC 6580
QY 2065 TGTTAATGCGAGTCTAGCAGAAAGAGATAGTAATTAAGTCTGCCAATTTGACAGACA 2124
Db 6581 TGTTAATGCGAGTCTAGCAGAAAGAGATAGTAATTAAGTCTGCCAATTTGACAGACA 6640
QY 2125 ATGCTAAACCATTAATGTAACAGCTGACATCTGTGAATTAATTTGTAACAGACCA 2184
Db 6641 ATGCTAAACCATTAATGTAACAGCTGACATCTGTGAATTAATTTGTAACAGACCA 6700
QY 2185 ACAACATATCAAGAAAAATGATCCGTATCCAGAGGGACACAGGAGGCAATTTGTTCAA 2244
Db 6701 ACAACATATCAAGAAAAATGATCCGTATCCAGAGGGACACAGGAGGCAATTTGTTCAA 6760
QY 2245 TAGAAAAAATAGAAAAATATGAGACAGACATTTGTAACATTAAGTACGAAAAATGGAATG 2304
Db 6761 TAGAAAAAATAGAAAAATATGAGACAGACATTTGTAACATTAAGTACGAAAAATGGAATG 6820
QY 2305 CCACTTTAAAAACAGATGCTAGCAAAATTAAGAGACAAATTTGGAATTAATTAACAAATTA 2364
Db 6821 ACACTTTAAAAACAGATGCTAGCAAAATTAAGAGACAAATTTGGAATTAATTAACAAATTA 6880
QY 2365 TCTTTAAGCAATCTCTCGAGAGGGACCCAGAAATTTGAAGCAAGTTTAATTTGAGAG 2424
Db 6881 TCTTTAAGCAATCTCTCGAGAGGGACCCAGAAATTTGAAGCAAGTTTAATTTGAGAG 6940
QY 2425 GGGAAATTTTCTACTGTAATTTCAACAACTGTTTATTAAGTACTGTTTAAATAGTACTT 2484
Db 6941 GGGAAATTTTCTACTGTAATTTCAACAACTGTTTATTAAGTACTGTTTAAATAGTACTT 7000
QY 2485 GGAATCTGAAGGCTCAATTAACATGTAAGGAAGTGAACAAATCACTCCATGACGAA 2544
|||||

Db 7001 GGAATCTGAAGGCTCAATTAACATGTAAGGAAGTGAACAAATCACTCCATGACGAA 7060
QY 2545 TAAACAAATTTAATTAACATGTGCGACAGATAGGAAGAAAGCAATGTATGCCCTCCATCA 2604
Db 7061 TAAACAAATTTAATTAACATGTGCGACAGATAGGAAGAAAGCAATGTATGCCCTCCATCA 7120
QY 2605 GTGGACAAATTAAGATGTTTCATCAAAATTAATTAAGTGGCTGTATTTAACAAGAGATGTGTGA 2664
Db 7121 GCGGACAAATTAAGATGTTTCATCAAAATTAATTAACAGGCTGTATTTAACAAGAGATGTGTGA 7180
QY 2665 ATTAACAACATGAGGTCCGAGATCTTCAGACCTTGAGAGAGCCGATTAAGAGCAATTTGA 2724
Db 7181 ATTAACAACATGAGGTCCGAGATCTTCAGACCTTGAGAGAGCCGATTAAGAGCAATTTGA 7240
QY 2725 GAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2784
Db 7241 GAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7300
QY 2785 AGGCAAAAGAAAGATGTGTCAGAGAGAAAAAAGACAGTGGGAATAGAGCTTTGTTCC 2844
Db 7301 AGGCAAAAGAAAGATGTGTCAGAGAGAAAAAAGACAGTGGGAATAGAGCTTTGTTCC 7360
QY 2845 TTGGGTTCTTTGGGACAGCAGAGAAAGCACTATGGGCTGACAGTCAATGAGCTACGCTAC 2904
Db 7361 TTGGGTTCTTTGGGACAGCAGAGAAAGCACTATGGGCTGACAGTCAATGAGCTACGCTAC 7420
QY 2905 AGGCCAGACAAATTAATGTCATATATGTCAGAGCAGAGCAATTTGCTGAGGCTATTG 2964
Db 7421 AGGCCAGACAAATTAATGTCATATATGTCAGAGCAGAGCAATTTGCTGAGGCTATTG 7480
QY 2965 AGGCCAGACAAATTAATGTCATATATGTCAGAGCAGAGCAATTTGCTGAGGCTATTG 3024
Db 7481 AGGCCAGACAAATTAATGTCATATATGTCAGAGCAGAGCAATTTGCTGAGGCTATTG 7540
QY 3025 TCTGCTGCTGGGAAAGTACTTAAGATCAACAGCTCTGGGGATTTGGGCTGTGCTG 3084
Db 7541 TCTGCTGCTGGGAAAGTACTTAAGATCAACAGCTCTGGGGATTTGGGCTGTGCTG 7600
QY 3085 GAAACATCAATTTGACACACTGCTGTGCTGTAAGTCTAGTGAATTAATTAATTTGCTG 3144
Db 7601 GAAACATCAATTTGACACACTGCTGTGCTGTAAGTCTAGTGAATTAATTAATTTGCTG 7660
QY 3145 AACAGATTTGGAATTAACATGACCTGGATGAGTGGGACAGAGAAATTAACATTTACACAA 3204
Db 7661 AACAGATTTGGAATTAACATGACCTGGATGAGTGGGACAGAGAAATTAACATTTACACAA 7720
QY 3205 GCTTAATACACTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3264
Db 7721 GCTTAATACACTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7780
QY 3265 TATTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3324
Db 7781 TATTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7840
QY 3325 GGTATATAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3384
Db 7841 GGTATATAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7900
QY 3385 CTGTACTTCTATATGAATAAGATTAAGGAGGATTTTACCATTAATGCTTACAGACC 3444
Db 7901 CTGTACTTCTATATGAATAAGATTAAGGAGGATTTTACCATTAATGCTTACAGACC 7960
QY 3445 ACCTCCCAATCCCGAGGGAGCCGACAGCCGCAAGAAATTAAGAAAGAGTGGAGAGA 3504
Db 7961 ACCTCCCAATCCCGAGGGAGCCGACAGCCGCAAGAAATTAAGAAAGAGTGGAGAGA 8020
QY 3505 GAGACAGACAGATCCATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3564
Db 8021 GAGACAGACAGATCCATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8080
QY 3565 TGGGAGGCTGTGCTCTTACAGTACCAAGGCTTGAAGAGCTTAATTAATTAATTAATTAAT 3624
Db 8081 TGGGAGGCTGTGCTCTTACAGTACCAAGGCTTGAAGAGCTTAATTAATTAATTAATTAAT 8140
|||||

Query Match	Best Local Similarity	Matches 2637: Conservative	Score 2598.2: 98.4%	DB 7: 0: Mismatches 38: Indels 6: Gaps 1:	Length 3156:
CC	(AAP60131) are claimed. The vector is pref. a member of the pV/env				
CC	family, e.g. pEV1.2 or 3/env 44-640 or 205-640.				
CC	(Updated on 25-MAR-2003 to correct PA field.)				
CC	(updated on 25-MAR-2003 to correct PI field.)				
xx					
sq	Sequence 3156 BP: 1095 A; 536 C; 765 G; 760 T; 0 other;				
QY	1291	CCATGAGGTGAAGGAGATGATGACGCTGTGGAGATGGGGGTGAATGGGGACCA	1350		
Db	476	CAATGAGGTGAAGGAGATGATGACGCTGTGGAGATGGGGGTGAATGGGGACCA	535		
QY	1351	TGCTCCTTGGGATATTGATGATCTGTAGTGTACAGAAAATTTGGGTCAACGCTATT	1410		
Db	536	TGCTCCTTGGGATATTGATGATCTGTAGTGTACAGAAAATTTGGGTCAACGCTATT	595		
QY	1411	ATGGGGTACCTGTGTGGAGAGACACACCACCTCTATTTTGTGCATCAGATGCTAAG	1470		
Db	596	ATGGGGTACCTGTGTGGAGAGACACACCACCTCTATTTTGTGCATCAGATGCTAAG	655		
QY	1471	CATATGATACAGAGGTACATATATGTTTGGGGCACACATGCTGTACCCACAGACCCA	1530		
Db	656	CATATGATACAGAGGTACATATATGTTTGGGGCACACATGCTGTACCCACAGACCCA	715		
QY	1531	ACCCACAGAAGTAGTATGTTGGTAAATGTGACAGAAAATTTTAAATGTGAAAAATGCA	1590		
Db	716	ACCCACAGAAGTAGTATGTTGGTAAATGTGACAGAAAATTTTAAATGTGAAAAATGCA	775		
QY	1591	TGTTAGAACAGATGATGAGATATATCACTTATGGATCAAAAGCCTAAAGCCATGTG	1650		
Db	776	TGTTAGAACAGATGATGAGATATATCACTTATGGATCAAAAGCCTAAAGCCATGTG	835		
QY	1651	TAAATTTAACCCCACTCTGTGTGTTTAAAGTCGACAGATTGAAGAATGATCTATA	1710		
Db	836	TAAATTTAACCCCACTCTGTGTGTTTAAAGTCGACAGATTGAAGAATGATCTATA	895		
QY	1711	CCAATAGTAGAGGGGAGAAATGATTAATGAGAGAAAGAGAGATAAAAAAGCTCTTTCA	1770		
Db	896	CCAATAGTAGAGGGGAGAAATGATTAATGAGAGAAAGAGAGATAAAAAAGCTCTTTCA	955		
QY	1771	ATATCAGACAGCATTAAGAGATTAAGGTGCGAAGAAATATATGATCTTTTATTAACCTG	1830		
Db	956	ATATCAGACAGCATTAAGAGATTAAGGTGCGAAGAAATATATGATCTTTTATTAACCTG	1015		
QY	1831	ATATAGTACCAATATGATTA-----TACACCTATAGGTGATTAAGTGTAAACCTGAG	1884		
Db	1016	ATATATATACCAATATATATATATATATCTACCTACCTATAGCTTACAAAGTGTAAACCTGAG	1075		
QY	1885	TCATTTACACAGAGCGCTGTCCAAAGGTATCCTTTGAGCCAAATCCCATACATTATTTGGCC	1944		
Db	1076	TCATTTACACAGAGCGCTGTCCAAAGGTATCCTTTGAGCCAAATCCCATACATTATTTGGCC	1135		
QY	1945	CGCGTGTGTTTGGCATTTCTAAATGTATAATTAAGACGTTCAATGTGAACAGAGACATGTA	2004		
Db	1136	CGCGTGTGTTTGGCATTTCTAAATGTATAATTAAGACGTTCAATGTGAACAGAGACATGTA	1195		
QY	2005	CAATGTGTACAGACATACATGTACACATGCAATCAGGCAGTAGTATCAACTCAACTGC	2064		
Db	1196	CAATGTGTACAGACATACATGTACACATGCAATCAGGCAGTAGTATCAACTCAACTGC	1255		
QY	2065	TGTTTAAATGGAGCTGTACAGAGAGAGATGTAATTAAGATTCGCGCAATTTCCACAGACA	2124		
Db	1256	TGTTTAAATGGAGCTGTGTACAGAGAGAGATGTAATTAAGATTCGCTCAATTTCCAGGACA	1315		
QY	2125	ATGCTTAAACCATATATAGTACAGCGTGAACACATCTGTGAATTAATTTGTACAAGACCA	2184		
Db	1316	ATGCTTAAACCATATATAGTACAGCGTGAACACATCTGTGAATTAATTTGTACAAGACCA	1375		
QY	2185	ACACACATACAGAAAAAGTATCCGCTATCCAGAGGGGACACAGGAGACATTTGTTACAA	2244		

```

Db 1376 ACAACATACAGAAAAAAATCCGTATCCAGAGGGGACGAGGAGACATTTGTTACAA 1435
Oy 2245 TAGGAAAAATAGAAATATGAGACAGACATTTGTAACATTAGTAGACAAATGGAATG 2304
Db 1436 TAGGAAAAATAGAAATATGAGACAGACATTTGTAACATTAGTAGACAAATGGAATG 1495
Oy 2305 CCACTTTAAAACAGATAGCTAGCAAAATTTAGAGAACAAATTTGAAATTAATAAACAATTA 2364
Db 1496 CCACCTTTAAAACAGATAGCTAGCAAAATTTAGAGAACAAATTTGAAATTAATAAACAATTA 1555
Oy 2365 TCTTTAAGCAATCTCTAGAGAGGGACCCAGAAATTTGTAGCCACAGTTTAAATTTGGAG 2424
Db 1556 TCTTTAAGCAATCTCTAGAGAGGGACCCAGAAATTTGTAGCCACAGTTTAAATTTGGAG 1615
Oy 2425 GGGATTTTCTAGCTGTAATTCACACAACTGTTTAATAGTACTGTTTAAATGTAATG 2484
Db 1616 GGGATTTTCTAGCTGTAATTCACACAACTGTTTAATAGTACTGTTTAAATGTAATG 1675
Oy 2485 GGAGTACTGAAGGGTCAAAATTAACACTGAAGAGTGAACAAATCCACTCCCATGCAAGAA 2544
Db 1676 GGAGTACTGAAGGGTCAAAATTAACACTGAAGAGTGAACAAATCCACTCCCATGCAAGAA 1735
Oy 2545 TAAACACATTTATTAACATGTGGCAGAGAGTAGGAAAAACAATGTATCCCTCCCATCA 2604
Db 1736 TAAACACATTTATTAACATGTGGCAGAGAGTAGGAAAAACAATGTATCCCTCCCATCA 1795
Oy 2605 GTGACACAAATAGATGTTCAATCAATATTTACTGGGCTCTATTAACAAGAGATGGTGSTA 2664
Db 1796 GCGGCAAAATTAAGATGTTCAATCAATATTTACTGGGCTCTATTAACAAGAGATGGTGSTA 1855
Oy 2665 ATAACACAAATGGTCCGAGATCTTCAGACCTGGAGAGCGGCAATGTAGGCAATTTGGA 2724
Db 1856 ATAACACAAATGGTCCGAGATCTTCAGACCTGGAGAGCGGCAATGTAGGCAATTTGGA 1915
Oy 2725 GAAGTGAATTTATTAATATTAAGTAGTAAAAATTTGAACCAATTTAGAGTACACCCACCA 2784
Db 1916 GAAGTGAATTTATTAATATTAAGTAGTAAAAATTTGAACCAATTTAGAGTACACCCACCA 1975
Oy 2785 AGGCAAGAGAGAGAGTGGTGAGAGAGAAAAAGAGCAGTGGGAATAGAGCTTTGTCTCC 2844
Db 1976 AGGCAAGAGAGAGAGTGGTGAGAGAGAAAAAGAGCAGTGGGAATAGAGCTTTGTCTCC 2035
Oy 2845 TTGGGCTTCTGGGAGCAGCAGAGAGACCTATGGGCTGACGTCAATGACGCTGACGCTAC 2904
Db 2036 TTGGGCTTCTGGGAGCAGCAGAGAGACCTATGGGCTGACGTCAATGACGCTGACGCTAC 2095
Oy 2905 AGGCGACAGCAATTTGTCATATAGTATAGTGCAGAGCAGACAAATTTGCTGAAGGCTATTTG 2964
Db 2096 AGGCGACAGCAATTTGTCATATAGTATAGTGCAGAGCAGACAAATTTGCTGAAGGCTATTTG 2155
Oy 2965 AGGCGCAACAGCATCTGTTGCAACTCAGAGTCTGGGGCATCAACAGCTCCAGGCAAGAA 3024
Db 2156 AGGCGCAACAGCATCTGTTGCAACTCAGAGTCTGGGGCATCAACAGCTCCAGGCAAGAA 2215
Oy 3025 TCCCTGGCTGGGAAAGATACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTCCTG 3084
Db 2216 TCCCTGGCTGGGAAAGATACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTCCTG 2275
Oy 3085 GAAACATCATTTGACACACTGCTGCTTGAATGCTAGTGTGAGTAATTAATCTCTGG 3144
Db 2276 GAAACATCATTTGACACACTGCTGCTTGAATGCTAGTGTGAGTAATTAATCTCTGG 2335
Oy 3145 AACAGATTTGGAATTAAGTAGACTGATGAGTGGGACAGAGAAATTAACAAATTAACAA 3204
Db 2336 AACAGATTTGGAATTAAGTAGACTGATGAGTGGGACAGAGAAATTAACAAATTAACAA 2395
Oy 3205 GCTTAATTAACACTCTCTAATTAAGATTCGCAAAACAGCAAGAAAGATCAACAGAAAT 3264
Db 2396 GCTTAATTAACACTCTCTAATTAAGATTCGCAAAACAGCAAGAAAGATCAACAGAAAT 2455
Oy 3265 TATTTGAATTTAGTAAATGGGCAAGTTTGGAAATTTGTTAAACATAACAAATTTGGCTGT 3324
Db 2456 TATTTGAATTTAGTAAATGGGCAAGTTTGGAAATTTGTTAAACATAACAAATTTGGCTGT 2515

```

```

Oy 3325 GGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3384
Db 2516 GGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2575
Oy 3385 CTGTACTTTCTAATAGTAAATAGAGTGTAGGAGGAGATTTACCAATTTATGCTTTAGACCC 3444
Db 2576 CTGTACTTTCTAATAGTAAATAGAGTGTAGGAGGAGATTTACCAATTTATGCTTTAGACCC 2635
Oy 3445 ACCTCCCAATCCCGAGGGGACCCGACAGGCCCGAAGGAATTAAGAAAGAGTGTAGAGA 3504
Db 2636 ACCTCCCAATCCCGAGGGGACCCGACAGGCCCGAAGGAATTAAGAAAGAGTGTAGAGA 2695
Oy 3505 GAGACAGAGACAGATCCATTTGATTAAGAACAGATCCCTTATGCACTTATCTGGAGACATC 3564
Db 2696 GAGACAGAGACAGATCCATTTGATTAAGAACAGATCCCTTATGCACTTATCTGGAGACATC 2755
Oy 3565 TGGGAGGCTGTGCTCTTACGCTACACACCGCTTGAGAGACTTACTTGTATTTGAAGA 3624
Db 2756 TGGGAGGCTGTGCTCTTACGCTACACACCGCTTGAGAGACTTACTTGTATTTGAAGA 2815
Oy 3625 GGATTTGGAACCTTCTGGAGAGCAGGGGCTGGGAAGCCCTCAATTTATGTTGGAATCTCC 3684
Db 2816 GGATTTGGAACCTTCTGGAGAGCAGGGGCTGGGAAGCCCTCAATTTATGTTGGAATCTCC 2875
Oy 3685 TACAGTATTTGAGTCAAGAACATTAAGAAATAGTGTGTTAACTGCTCAATGCCACAGCA 3744
Db 2876 TACAGTATTTGAGTCAAGAACATTAAGAAATAGTGTGTTAACTGCTCAATGCCACAGCA 2935
Oy 3745 TAGCAGTAGTCTGAGGGGACAGATAGGCTTATAGAAATTTTCAACAGCAGCTTATAGACCTA 3804
Db 2936 TAGCAGTAGTCTGAGGGGACAGATAGGCTTATAGAAATTTTCAACAGAACTTATAGACCTA 2995
Oy 3805 TTGCGCACATCTCTAGAAAGATTAAGACAGGCGCTTGGAAGGATTTGCTATTAAGATGGGT 3864
Db 2996 TTGCGCACATCTCTAGAAAGATTAAGACAGGCGCTTGGAAGGATTTGCTATTAAGATGGGT 3055
Oy 3865 GGCAGATGCTCAAAAAGTAGTGTGATTTGATGGCTGCTGTAAAGGAAAGATAGAGAGA 3924
Db 3056 GGCAGATGCTCAAAAAGTAGTGTGATTTGATGGCTGCTGTAAAGGAAAGATAGAGAGA 3115
Oy 3925 GCTGAGCCAGCAGCAGATGGGCTGGAGCAGTATCTGAGAGA 3965
Db 3116 GCTGAGCCAGCAGCAGATGGGCTGGAGCAGTATCTGAGAGA 3156

RESULT 15
AAH76385
ID AAH76385 standard; DNA; 8932 BP.
XX
AC AAH76385;
XX
AC 30-NOV-2001 (first entry)
XX
DE Nucleotide sequence of HIV-1 isolate BH10.
XX
KW HIV-1; gp120; BH10; vaccine; immunization; ds.
XX
OS Human immunodeficiency virus type 1.
XX
PN US6268484-B1.
XX
XX 31-JUL-2001.
XX
PD 30-JUL-1998; 98US-0124900.
XX
PE 07-JUN-1995; 95US-0478536.
XX
PR 19-APR-1995; 95WO-EP01481.
XX
PA (POLY-) POLYMUN SCI IMMUNOBIOLOGISCHE FORSCHUNG.
PI Kätlinger H, Buchacher A, Ernst W, Ballaun C, Purtscher M;
PI Trkola A, Predl R, Schmatz C, Klima A, Steindl F, Muster T;

```


QY 3085 GAAACCTATTGTCACCACTGCTGTGCTTGAATGCTAGTGAATTAATCTCTGG 3144
|||||
Db 7378 GAAACTCATTTGTCACCACTGCTGTGCTTGAATGCTAGTGAATTAATCTCTGG 7437
QY 3145 AACAGATTGGAAATTAACATGACCTGTGAGTGGGACAGAAATTAACATTAACAA 3204
|||||
Db 7438 AACAGATTGGAAATTAACATGACCTGTGAGTGGGACAGAAATTAACATTAACAA 7497
QY 3205 GCTTAATACCTCTTATTAATGAAGATCCGAAAAACGCAAGAAATGAACAGAAAT 3264
|||||
Db 7498 GCTTAATACCTCTTATTAATGAAGATCCGAAAAACGCAAGAAATGAACAGAAAT 7557
QY 3265 TATTGAATTAGATTAATGGCAAGTTTGTGAATTGGTTAAATTAACAAATGGCTGT 3324
|||||
Db 7558 TATTGAATTAGATTAATGGCAAGTTTGTGAATTGGTTAAATTAACAAATGGCTGT 7617
QY 3325 GGTATATATAATTAATCATATATAGTAGAGGCTGTAGGTTAAGAAATAGTTTGG 3384
|||||
Db 7618 GGTATATATAATTAATCATATATAGTAGAGGCTGTAGGTTAAGAAATAGTTTGG 7677
QY 3385 CTGTACTTCTATAGTAATAGATTAAGGATATTCACATTTATCTTCAGACCC 3444
|||||
Db 7678 CTGTACTTCTATAGTAATAGATTAAGGATATTCACATTTATCTTCAGACCC 7737
QY 3445 ACCTCCCATCCGAGGGGACCCGACAGGCCGGAAGAAATAGAAGAGGTGAGAGA 3504
|||||
Db 7738 ACCTCCCATCCGAGGGGACCCGACAGGCCGGAAGAAATAGAAGAGGTGAGAGA 7797
QY 3505 GAGACAGAGACAGATCCATTCGATTAAGTGAACGATCCTTACGACTTATCTGGAGCATC 3564
|||||
Db 7798 GAGACAGAGACAGATCCATTCGATTAAGTGAACGATCCTTACGACTTATCTGGAGCATC 7857
QY 3565 TGGGAGGCCGTGGCTCTTACGATACCAAGCTTGAAGACTTACTCTGATTTAGAGA 3624
|||||
Db 7858 TGGGAGGCCGTGGCTCTTACGATACCAAGCTTGAAGACTTACTCTGATTTAGAGA 7917
QY 3625 GGATTTGTGAACCTTGTGGAGCAGAGGGGTGGGAAAGCCCTCAATATTTGGTGAATCTCC 3684
|||||
Db 7918 GGATTTGTGAACCTTGTGGAGCAGAGGGGTGGGAAAGCCCTCAATATTTGGTGAATCTCC 7977
QY 3685 TACAGTATTGGAGTCAAGAACTTAAGAAATAGTCTTTAACTTGCTCAATGCCACAGCA 3744
|||||
Db 7978 TACAGTATTGGAGTCAAGAACTTAAGAAATAGTCTTTAACTTGCTCAATGCCACAGCTA 8037
QY 3745 TACAGTATTGGAGTCAAGAACTTAAGAAATAGTCTTTAACTTGCTCAATGCCACAGCTA 3804
|||||
Db 8038 TACAGTATTGGAGTCAAGAACTTAAGAAATAGTCTTTAACTTGCTCAATGCCACAGCTA 8097
QY 3805 TTGGCCACATACCTAGAAATTAAGACAGGGCTTGAAGGATTTTCTATAAGATGGGT 3864
|||||
Db 8098 TTGGCCACATACCTAGAAATTAAGACAGGGCTTGAAGGATTTTCTATAAGATGGGT 8157
QY 3865 GGCAGTGTGCACAAAAGTACTGTGATTTGATGGCTGCTTAAGGGAAGAAATGAGAGCA 3924
|||||
Db 8158 GGCAGTGTGCACAAAAGTACTGTGATTTGATGGCTGCTTAAGGGAAGAAATGAGAGCA 8217
QY 3925 GCTGAGCCAGCAGATGGGTTGGGAGCAGTATCTGAGATCTAGA 3971
|||||
Db 8218 GCTGAGCCAGCAGATGGGTTGGGAGCAGTATCTGAGATCTAGA 8264

Search completed: September 17, 2003, 08:05:26
Job time : 1488 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2003, 07:40:41 : Search time 320 Seconds
(without alignments) 8591.799 Million cell updates/sec

Title: US-09-913-159A-10

Perfect score: 6229

Sequence: 1 ctgacgcgcctctgacgcgc.....attcccccgaagtgcac 6229

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents, NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCYUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2676.2	43.0	9709	2	US-08-188-583-5
2	2676.2	43.0	9709	3	US-08-388-353-1
3	2676.2	43.0	9709	4	US-08-488-551B-1
4	2676.2	43.0	9709	4	US-09-309-572-15
5	2636.6	42.3	15581	3	US-08-646-538-35
6	2636.6	42.3	15581	3	US-09-503-222-35
7	2597.8	41.7	8933	3	US-09-124-900-1
8	2597.8	41.7	8933	3	US-08-463-210-4
9	2597.8	41.7	8933	3	US-09-620-958A-3
10	2597.8	41.7	8933	3	US-09-620-958A-4
11	2597.8	41.7	8933	3	US-09-620-958A-9
12	2593	41.6	9719	4	US-09-700-304-1
13	2590.8	41.6	9745	6	5462872-1
14	2558.4	41.1	12479	4	US-09-318-138-13
15	2558.4	41.1	12494	3	US-08-935-312-13
16	2558.4	41.1	12494	3	US-08-848-760B-33
17	2554.6	41.0	3563	3	US-08-463-210-6
18	2486.6	39.9	2571	2	US-07-916-098A-1
19	2411.6	38.7	2644	3	US-08-472-240A-9
20	2303.4	37.0	2531	3	US-07-956-483-18
21	2205.6	35.4	3807	1	US-08-022-835-5
22	2205.6	35.4	3807	1	US-08-388-809-5
23	2205.6	35.4	3807	2	US-08-647-714-5
24	2199.4	35.3	2730	3	US-08-728-122-1
25	2195.4	35.2	2696	4	US-09-325-131B-1
26	2141.6	34.4	9737	2	US-08-944-449-7
27	2141.6	34.4	9737	4	US-09-353-362-7

28	2118.8	34.0	2552	2	US-08-448-603A-27	Sequence 27, Appl
29	2118.8	34.0	2552	3	US-09-134-075-27	Sequence 27, Appl
30	2118.8	34.0	2552	4	US-09-492-739-27	Sequence 27, Appl
31	2113	33.9	9746	1	US-08-022-835-3	Sequence 3, Appl
32	2113	33.9	9746	2	US-08-388-809-3	Sequence 3, Appl
33	2113	33.9	9746	2	US-08-647-714-3	Sequence 3, Appl
34	2111.8	33.9	4527	2	US-08-944-449-8	Sequence 8, Appl
35	2111.8	33.9	4527	4	US-09-353-362-8	Sequence 8, Appl
36	2101.8	33.7	9739	1	US-08-022-835-1	Sequence 1, Appl
37	2101.8	33.7	9739	2	US-08-388-809-1	Sequence 1, Appl
38	2101.8	33.7	9739	2	US-08-647-714-1	Sequence 1, Appl
39	2091.2	33.6	6474	3	US-08-358-928-66	Sequence 66, Appl
40	2091.2	33.6	6474	3	US-08-358-928-66	Sequence 66, Appl
41	2090.6	33.6	6926	3	US-08-651-472-69	Sequence 69, Appl
42	2090.6	33.6	6926	3	US-08-358-928-69	Sequence 69, Appl
43	2088.2	33.5	2573	2	US-08-448-603A-29	Sequence 29, Appl
44	2088.2	33.5	2573	3	US-09-134-075-29	Sequence 29, Appl
45	2088.2	33.5	2573	4	US-09-492-739-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-08-188-583-5
Sequence 5, Application US/08188583

Patent No. 5851813

GENERAL INFORMATION:

APPLICANT: Destroiers, Ronald C.

TITLE OF INVENTION: PRIMATE LENTIVIRUS VACCINES

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 50Z or 55SX

OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)

SOFTWARE: WordPerfect (Version 5.0)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/188, 583

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/27,494

FILING DATE: July 9, 1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/551,945

FILING DATE: July 12, 1990

ATTORNEY/AGENT INFORMATION:

NAME: Freeman, John W.

REGISTRATION NUMBER: Reg. No. 5851813 29,066

REFERENCE/DOCKET NUMBER: 00246/079002

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 9709

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-188-583-5

Query Match 43.0%, Score 2676.2, DB 2, Length 9709;

Best Local Similarity 99.9%, Pred. No. 0;

Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1291 CCATGAGAGTGAAGAGACATATACGACTTGTGAGATGGGGGTGAAATGGGGACCA 1350
| | | | |
Db 6219 CAATGAGAGTGAAGAGACATATACGACTTGTGAGATGGGGGTGAAATGGGGACCA 6278
QY 1351 TGCCTCTGGGATATGTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1410
| | | | |
Db 6279 TGCCTCTGGGATATGTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6338
QY 1411 ATGGGGTACCTGT 1470
| | | | |
Db 6339 ATGGGGTACCTGT 6398
QY 1471 CATATGATACGAGGTACATATATTTTGGGCCACACATGCTGTGTGTGTGTGTGTGTGT 1530
| | | | |
Db 6399 CATATGATACGAGGTACATATATTTTGGGCCACACATGCTGTGTGTGTGTGTGTGTGT 6458
QY 1531 ACCCACAAGAGT 1590
| | | | |
Db 6459 ACCCACAAGAGT 6518
QY 1591 TGGTGAACACAGATCATGAGGATATATCACTTTATGCGATCAAGCCCTAAAGCCATGTG 1650
| | | | |
Db 6519 TGGTGAACACAGATCATGAGGATATATCACTTTATGCGATCAAGCCCTAAAGCCATGTG 6578
QY 1651 TAAATTAACCCCACTGT 1710
| | | | |
Db 6579 TAAATTAACCCCACTGT 6638
QY 1711 CCATAGTAGTAGGAGGAGATGATTAATGAGAAAGAGATTAATAAATGCTCTTTCA 1770
| | | | |
Db 6639 CCATAGTAGTAGGAGGAGATGATTAATGAGAAAGAGATTAATAAATGCTCTTTCA 6698
QY 1771 ATATCAGACAAGATGAAGATTAAGTGCAGAAAGAAATATGCAATCTTTTATTAATCTTG 1830
| | | | |
Db 6699 ATATCAGACAAGATGAAGATTAAGTGCAGAAAGAAATATGCAATCTTTTATTAATCTTG 6758
QY 1831 ATATAGTAGCAATAGATATATCCAGCTATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1890
| | | | |
Db 6759 ATATAGTAGCAATAGATATATCCAGCTATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6818
QY 1891 CACAGGCTGTCCAAAGGATCTTTGAGCCATTTCCCATATATATGTCGCCGGGTG 1950
| | | | |
Db 6819 CACAGGCTGTCCAAAGGATCTTTGAGCCATTTCCCATATATATGTCGCCGGGTG 6878
QY 1951 GTTTTGGGATCTAAATGTATATATTAAGAGCTTCAATGAGAAAGACCATGTCAAAATG 2010
| | | | |
Db 6879 GTTTTGGGATCTAAATGTATATATTAAGAGCTTCAATGAGAAAGACCATGTCAAAATG 6938
QY 2011 TCACACAGATGATGTATACATGTGAATCAGGCCAGTATGTCAACTCACTGCTTTAA 2070
| | | | |
Db 6939 TCACACAGATGATGTATACATGTGAATCAGGCCAGTATGTCAACTCACTGCTTTAA 6998
QY 2071 ATGCGATCTAGCAGAGAGATGTAGTAATTTAGATGCGCAATTTTCACAGACATGCTA 2130
| | | | |
Db 6999 ATGCGATCTAGCAGAGAGATGTAGTAATTTAGATGCGCAATTTTCACAGACATGCTA 7058
QY 2131 AAACCATTAATAGTACAGCTGAACACATCTGTGAATTAATTTGTACAAAGACCAACAACA 2190
| | | | |
Db 7059 AAACCATTAATAGTACAGCTGAACACATCTGTGAATTAATTTGTACAAAGACCAACAACA 7118
QY 2191 ATACAGAAAAAGTATCCGATTCAGAGGGGACCAAGGAGACATTTGTGTACAAATAGGAA 2250
| | | | |
Db 7119 ATACAGAAAAAGTATCCGATTCAGAGGGGACCAAGGAGACATTTGTGTACAAATAGGAA 7178
QY 2251 AAATAGGAAATATAGCAACGACATGTATACATTAGAGGCAAAATGAAATGCCACTT 2310
| | | | |
Db 7179 AAATAGGAAATATAGCAACGACATGTATACATTAGAGGCAAAATGAAATGCCACTT 7238
QY 2311 TAAACAGATAGTACGAAATTAAGAGAAATTTGGAAATTAATTAACCAATTAATCTTTA 2370
| | | | |
Db 7239 TAAACAGATAGTACGAAATTAAGAGAAATTTGGAAATTAATTAACCAATTAATCTTTA 7298
QY 2371 AGCAATCTCTAGAGGGGGACCAAGAAATTTGTACGACAGTTTAAATTTGTGAGGGGAAAT 2430
| | | | |

Db 7299 AGCAATCTCTAGAGGGGGACCAAGAAATTTGTACGACAGTTTAAATTTGTGAGGGGAAAT 7358
QY 2431 TTTTCTACTGTAAATTCACACACACATGTTTAAATAGTGTGTGTGTGTGTGTGTGTGTGT 2490
| | | | |
Db 7359 TTTTCTACTGTAAATTCACACACACATGTTTAAATAGTGTGTGTGTGTGTGTGTGTGTGT 7418
QY 2491 CTGAAGGGTCAAAATTAACACTGAAGAGTGACACAAATCCACTCCAGAGAAATTAAC 2550
| | | | |
Db 7419 CTGAAGGGTCAAAATTAACACTGAAGAGTGACACAAATCCACTCCAGAGAAATTAAC 7478
QY 2551 AATTTATTAACATGTGGCAGGAATAGGAAAGCAATGTATGCCCCCTCCATAGTGTGAC 2610
| | | | |
Db 7479 AATTTATTAACATGTGGCAGGAATAGGAAAGCAATGTATGCCCCCTCCATAGTGTGAC 7538
QY 2611 AAATTAATGTTCATCAAAATATTTACTGGGCTGTATTAACAAAGATGTGTGTGTGTGTGT 2670
| | | | |
Db 7539 AAATTAATGTTCATCAAAATATTTACTGGGCTGTATTAACAAAGATGTGTGTGTGTGTGT 7598
QY 2671 ACAATGGGTCCGAGATCTTCAGACCTGGAGAGGCGATATGAGGACAAATTTGGAAGTG 2730
| | | | |
Db 7599 ACAATGGGTCCGAGATCTTCAGACCTGGAGAGGCGATATGAGGACAAATTTGGAAGTG 7658
QY 2731 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2790
| | | | |
Db 7659 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7718
QY 2791 AGAGAAAGT 2850
| | | | |
Db 7719 AGAGAAAGT 7778
QY 2851 TCTTGGGAGCAGCAGGAAGCCTATGAGCTGCACGTCAATGACGCTGACGCTACAGCCCA 2910
| | | | |
Db 7779 TCTTGGGAGCAGCAGGAAGCCTATGAGCTGCACGTCAATGACGCTGACGCTACAGCCCA 7838
QY 2911 GACAAATATGTCTGTATATATGTGAGCAGCAGCAATTTGCTGAGGCTATTTGAGGCGC 2970
| | | | |
Db 7839 GACAAATATGTCTGTATATATGTGAGCAGCAGCAATTTGCTGAGGCTATTTGAGGCGC 7898
QY 2971 AACGCAATCTGT 3030
| | | | |
Db 7899 AACGCAATCTGT 7958
QY 3031 CTGTGGAAGATTAACCTAAAGGATCAACAGCTCTGTGGGATTTGGGGTGTCTGTGAAATC 3090
| | | | |
Db 7959 CTGTGGAAGATTAACCTAAAGGATCAACAGCTCTGTGGGATTTGGGGTGTCTGTGAAATC 8018
QY 3091 TCATTTGCACACCTGT 3150
| | | | |
Db 8019 TCATTTGCACACCTGT 8078
QY 3151 TTTTGAATTAACATGACCTGT 3210
| | | | |
Db 8079 TTTTGAATTAACATGACCTGT 8138
QY 3211 TACACTCCTTAATTTGAAGATTCGAAACACGACAAAGAAAGATGAACAAATTAATTTG 3270
| | | | |
Db 8139 TACACTCCTTAATTTGAAGATTCGAAACACGACAAAGAAAGATGAACAAATTAATTTG 8198
QY 3271 AATTAATTAATTTGGGCAAGTTTGTGAAATTTGTTAACTATTAACAAATTTGGCTGTGTATA 3330
| | | | |
Db 8199 AATTAATTAATTTGGGCAAGTTTGTGAAATTTGTTAACTATTAACAAATTTGGCTGTGTATA 8258
QY 3331 TAAATTAATTTCAATATGATGTAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3390
| | | | |
Db 8259 TAAATTAATTTCAATATGATGTAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8318
QY 3391 TTTCTATAGTAATAGATTAGGAGGGAATTTCCACATTAATCTTTACAGCCCACTCC 3450
| | | | |
Db 8319 TTTCTATAGTAATAGATTAGGAGGGAATTTCCACATTAATCTTTACAGCCCACTCC 8378
QY 3451 CAATCCGAGGGGACCCGACAGGCCCGAAGGAATAGAGAAAGAGGTGAGAGAGAGACA 3510
| | | | |

Db 8379 CAATCCGAGGGGACCCGACAGGCGCGAAGNATGAAAGACAGGTGGAGAGAGACA 8438
QY 3511 GAGNAGATCCATTGGATTAGTGAACGGATCTTACGACTTATCTGGGACAGCTTCGGCA 3570
Db 8439 GAGACAGATCCATTGGATTAGTGAACGGATCTTACGACTTATCTGGGACAGCTTCGGCA 8498
QY 3571 GCCTGTGCTCTTTCAGTACCAACCGCTTGAGAGACTTCTCTGATTTGATTAACGAGGATTG 3630
Db 8499 GCCTGTGCTCTTTCAGTACCAACCGCTTGAGAGACTTCTCTGATTTGATTAACGAGGATTG 8558
QY 3631 TGGAACTTCTGGGACGACAGGGGGTGGGAAGCCCTCAAAATTTGATGGAATCTCTACAGT 3690
Db 8559 TGGAACTTCTGGGACGACAGGGGGTGGGAAGCCCTCAAAATTTGATGGAATCTCTACAGT 8618
QY 3691 ATTGAGTCTGAGAACTAAAGAAATAGTCTGTTAACTTCTCTCAATCCACAGCCATAGCAG 3750
Db 8619 ATTGAGTCTGAGAACTAAAGAAATAGTCTGTTAACTTCTCTCAATCCACAGCCATAGCAG 8678
QY 3751 TAGCTGAGGGGACAGATAGGGTTATAGAAGTATTCAAGCAGCTTATAGACTATTCGCC 3810
Db 8679 TAGCTGAGGGGACAGATAGGGTTATAGAAGTATTCAAGCAGCTTATAGACTATTCGCC 8738
QY 3811 ACATACCTAGAAAGATAAGACAGGGCTTGAAGGATTTTGTATATAGATGGGTGGCAAG 3870
Db 8739 ACATACCTAGAAAGATAAGACAGGGCTTGAAGGATTTTGTATATAGATGGGTGGCAAG 8798
QY 3871 TGGTAAAAAGTAGTGTATTTGATGGCTGCTGTGAAGGAAAGATAGACGAGCTGAG 3930
Db 8799 TGGTAAAAAGTAGTGTATTTGATGGCTGCTGTGAAGGAAAGATAGACGAGCTGAG 8858
QY 3931 CCAGCAGCAGATGGGGTGGAGCAGTATCTGAGATCTAGA 3971
Db 8859 CCAGCAGCAGATGGGGTGGAGCAGTATCTGAGACCTAGA 8899

RESULT 2
US-08-388-353-1

: Sequence 1, Application US/08388353

: Patent No. 6010895

: GENERAL INFORMATION:

: APPLICANT: Deacon, Nicholas J.

: APPLICANT: Learmont, Jennifer C.

: APPLICANT: McPhee, Dale A.

: APPLICANT: Crowe, Suzanne

: APPLICANT: Cooper, David

: TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1

: NUMBER OF SEQUENCES: 800

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Scully, Scott, Murphy & Presser

: STREET: 400 Garden City Plaza

: CITY: Garden City

: STATE: New York

: COUNTRY: United States

: ZIP: 11530

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patent Release #1.0, Version #1.25

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/388,353

: FILING DATE: 14-FEB-1995

: CLASSIFICATION: 424

: ATTORNEY/AGENT INFORMATION:

: NAME: Digilio, Frank S.

: REGISTRATION NUMBER: 31,346

: REFERENCE/DOCKET NUMBER: 9606

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (516) 742-4343

: TELEFAX: (516) 742-4366

: TELEX: 230 901 SANS UR

: INFORMATION FOR SEQ ID NO: 1:

: SEQUENCE CHARACTERISTICS:

: ;

LENGTH: 9709 base pairs
type: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-388-353-1

Query Match 43.0%; Score 2676.2; DB 3; Length 9709;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1291 CCATGAGGTGAGAGAGATGATGACGCTTGAGATGGGGGTGGAATGGGGCAGCA 1350
Db 6219 CAATGAGGTGAGAGAGATGATGACGCTTGAGATGGGGGTGGAATGGGGCAGCA 6278
QY 1351 TGTCTCTGGGATTTGATGATCTGATGCTGACAGAAAATTTGTGGGTCAACGCTATT 1410
Db 6279 TGTCTCTGGGATTTGATGATCTGATGCTGACAGAAAATTTGTGGGTCAACGCTATT 6338
QY 1411 ATGGGGTACCTGTGTGGAAGGAGCAACACCACTCTATTGTCATCAGATGCTAAG 1470
Db 6339 ATGGGGTACCTGTGTGGAAGGAGCAACACCACTCTATTGTCATCAGATGCTAAG 6398
QY 1471 CATATGATACAGAGGTACATATGTTGGGCCACATGCGGTGACCAAGCCCA 1530
Db 6399 CATATGATACAGAGGTACATATGTTGGGCCACATGCGGTGACCAAGCCCA 6458
QY 1531 ACCCAAGAGTACTATTGTTAATGTGACAGAAAATTTTAACTGTGAAAATGACA 1590
Db 6459 ACCCAAGAGTACTATTGTTAATGTGACAGAAAATTTTAACTGTGAAAATGACA 6518
QY 1591 TGGTGAACAGATGATGAGGATATATACGTTTATGGATCAAAAGCCTAAAGCATGTG 1650
Db 6519 TGGTGAACAGATGATGAGGATATATACGTTTATGGATCAAAAGCCTAAAGCATGTG 6578
QY 1651 TAAATTAACCCCACTGTGTAGTTTAAAGTGCATGATTTGAAGATGACTAATA 1710
Db 6579 TAAATTAACCCCACTGTGTAGTTTAAAGTGCATGATTTGAAGATGACTAATA 6638
QY 1711 CCAATAGTAGTAGCGGAGAAATGATTAATGAGAAAGAGAGATAAATGCTCTTCA 1770
Db 6639 CCAATAGTAGTAGCGGAGAAATGATTAATGAGAAAGAGAGATAAATGCTCTTCA 6698
QY 1771 ATATCAGCAAAAGCTAAGATATAGTGTGCAAGAAATATGCACTTCTTTATAACTTG 1830
Db 6699 ATATCAGCAAAAGCTAAGATATAGTGTGCAAGAAATATGCACTTCTTTATAACTTG 6758
QY 1831 ATATGATCAATATAGTATACACGCTATAGTTGATAGTTGTAACACCTCAGCATTA 1890
Db 6759 ATATGATCAATATAGTATACACGCTATAGTTGATAGTTGTAACACCTCAGCATTA 6818
QY 1891 CACAGGCTGTCCAAAGGTATCTTTGAGCAATTCACATATTTATGTCGCCGCGTG 1950
Db 6819 CACAGGCTGTCCAAAGGTATCTTTGAGCAATTCACATATTTATGTCGCCGCGTG 6878
QY 1951 GTTTTGGCATTTCTAAATGTAATATTAAGAGCTTCAATGGAACAGACCATGTACAAATG 2010
Db 6879 GTTTTGGCATTTCTAAATGTAATATTAAGAGCTTCAATGGAACAGACCATGTACAAATG 6938
QY 2011 TCAGCAGGTCAATATGATGATGGAATGGAAGCCAGCTGTATCAACTGCTGTAA 2070
Db 6939 TCAGCAGGTCAATATGATGGAATGGAAGCCAGCTGTATCAACTGCTGTAA 6998
QY 2071 ATGGCAGTCTAGCAGAGAGATGTAATTTAGATCTGCCAATTTTCACAGCAATGTGA 2130
Db 6999 ATGGCAGTCTAGCAGAGAGATGTAATTTAGATCTGCCAATTTTCACAGCAATGTGA 7058
QY 2131 AACCATATATGATCAGCTGGAACACATCTGTGAATTTAATTTGACAGACCAACACA 2190
Db 7059 AACCATATATGATCAGCTGGAACACATCTGTGAATTTAATTTGACAGACCAACACA 7118
QY 2191 ATACAGAAAAGATCCGTATCCAGAGGAGCAGGAGAGCATTTGTTCAATATAGAA 2250
Db 7118 ATACAGAAAAGATCCGTATCCAGAGGAGCAGGAGAGCATTTGTTCAATATAGAA 2250

Db 7119 ATACAGAAAAAGTATCCGATCCAGAGGGGACGAGGAGACATTGTTCAATAGGAA 7178
Qy 2251 AAATAGGAATATGAGACAGACATTGTATACATTAGTAGAGCAAAATGAGATCCACTT 2310
Db 7179 AAATAGGAATATGAGACAGACATTGTATACATTAGTAGAGCAAAATGAGATCCACTT 7238
Qy 2311 TAAACAGATATGCTAGCAAAATTAAGAGACAAATTTGGAATTAATTAACAAATATCTTTA 2370
Db 7239 TAAACAGATATGCTAGCAAAATTAAGAGACAAATTTGGAATTAATTAACAAATATCTTTA 7298
Qy 2371 AGCATCCTCAGAGGGGACCCAGAAATTTGAGCAGCAAGTTTAAATGTTGAGGGGAGAT 2430
Db 7299 AGCATCCTCAGAGGGGACCCAGAAATTTGAGCAGCAAGTTTAAATGTTGAGGGGAGAT 7358
Qy 2431 TTTTCTACTGTAAATCAACACACTGTTTAAATAGTACTTGTAAATAGTACTTGGACTA 2490
Db 7359 TTTTCTACTGTAAATCAACACACTGTTTAAATAGTACTTGTAAATAGTACTTGGACTA 7418
Qy 2491 CTGAAGGTCAAATTAACACTGAGAGAGTGCACACATCACACTCCCATGCAAAATAAAC 2550
Db 7419 CTGAAGGTCAAATTAACACTGAGAGAGTGCACACATCACACTCCCATGCAAAATAAAC 7478
Qy 2551 AATTATTAACATGTCGAGAGAGTAAAGGAAAGCAATGATGCCCCCTCCATCATGTCGAC 2610
Db 7479 AATTATTAACATGTCGAGAGAGTAAAGGAAAGCAATGATGCCCCCTCCATCATGTCGAC 7538
Qy 2611 AAATAGATGTTCAATCAAAATTTACTGGGCTGCTATTAAACAAGATGTTGTTAATAACA 2670
Db 7539 AAATAGATGTTCAATCAAAATTTACTGGGCTGCTATTAAACAAGATGTTGTTAATAACA 7598
Qy 2671 ACAATGGTCCGAGATCTTCAGACCTGAGAGAGCGGATATGAGGACAAATTTGAGAAATG 2730
Db 7599 ACAATGGTCCGAGATCTTCAGACCTGAGAGAGCGGATATGAGGACAAATTTGAGAAATG 7658
Qy 2731 AATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2790
Db 7659 AATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7718
Qy 2791 AGAGAGAGTGTGTCAG 2850
Db 7719 AGAGAGAGTGTGTCAG 7778
Qy 2851 TCTTGGAGAGCAG 2910
Db 7779 TCTTGGAGAGCAG 7838
Qy 2911 GACATTAATGTCGATATGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2970
Db 7839 GACATTAATGTCGATATGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7898
Qy 2971 AACAGCATGTCGTCACATCAGACTGAGGAGATCAACACAGCTCCAGGCAAGATTCCTG 3030
Db 7899 AACAGCATGTCGTCACATCAGACTGAGGAGATCAACACAGCTCCAGGCAAGATTCCTG 7958
Qy 3031 CTGTGGAAGATATCTTAAGAGATCAACAGCTCCGAGGAGATTTGGGCTGCTGGAAGC 3090
Db 7959 CTGTGGAAGATATCTTAAGAGATCAACAGCTCCGAGGAGATTTGGGCTGCTGGAAGC 8018
Qy 3091 TCATTTGACACACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3150
Db 8019 TCATTTGACACACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 8078
Qy 3151 TTTTGAATTAACATGACCTGATGAGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3210
Db 8079 TTTTGAATTAACATGACCTGATGAGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 8138
Qy 3211 TACACTCCTTAATTAAGAGATCGCAAAACAGCAAGAGAGAGAGAGAGAGAGAGAGAG 3270
Db 8139 TACACTCCTTAATTAAGAGATCGCAAAACAGCAAGAGAGAGAGAGAGAGAGAGAGAG 8198
Qy 3271 AATTAGATTAATGAGGAGAGTTTGTGGAATGCTTAACATTAACAAATTTGCTGTGATATA 3330
Db 8199 AATTAGATTAATGAGGAGAGTTTGTGGAATGCTTAACATTAACAAATTTGCTGTGATATA 8258

Qy 3331 TAAATTAATTCAT 3390
Db 8259 TAAATTAATTCAT 8318
Qy 3391 TTTTCTAT 3450
Db 8319 TTTTCTAT 8378
Qy 3451 CAATCCGAGAGGAG 3510
Db 8379 CAATCCGAGAGGAG 8438
Qy 3511 GAGACAGATCCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 3570
Db 8439 GAGACAGATCCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 8498
Qy 3571 GCCTGTGCTCTTACAGTACAGCCGCTTTGAGAGACTTACTTGTATATATATATATATATAT 3630
Db 8499 GCCTGTGCTCTTACAGTACAGCCGCTTTGAGAGACTTACTTGTATATATATATATATATAT 8558
Qy 3631 TGAACATTCCTGAG 3690
Db 8559 TGAACATTCCTGAG 8618
Qy 3691 ATTGAGTCAAGAACTAAAGATATAGTCTGTTAACTTCTTCAATGCCACAGCCCATAGCAG 3750
Db 8619 ATTGAGTCAAGAACTAAAGATATAGTCTGTTAACTTCTTCAATGCCACAGCCCATAGCAG 8678
Qy 3751 TAGCTGAGGAG 3810
Db 8679 TAGCTGAGGAG 8738
Qy 3811 ACATACCTAGAGAGATTAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3870
Db 8739 ACATACCTAGAGAGATTAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8798
Qy 3871 TGTCAAAAAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3930
Db 8799 TGTCAAAAAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8858
Qy 3931 CCAGCAGCAGATGGGCTGGGAGCAGTATCTGAGATCTAGA 3971
Db 8859 CCAGCAGCAGATGGGCTGGGAGCAGTATCTGAGATCTAGA 8899

RESULT 3
US-08-488-551B-1
; Sequence 1, Application US/08488551B
; Patent No. 6015661
; GENERAL INFORMATION:
; APPLICANT: Nicholas J. Deacon
; APPLICANT: Dale A. McPhee
; APPLICANT: David Cooper
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 841
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,551B
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PM3864 (AU)

FLING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FLING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FLING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FLING DATE: 14-FEB-1995
APPLICATION NUMBER: PM3021/95
FLING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO
REFERENCE/DOCKET NUMBER: 9606Z
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9709 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-488-551B-1

Query Match 43.0%; Score 2676.2; DB 3; Length 9709;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1291 CCATGAGAGTGAAGAGAGATATCAGCATTGTGGAGATGGGGTGGNAATGGGGCACC 1350
6219 CAATGAGAGTGAAGAGAGATATCAGCATTGTGGAGATGGGGGGAATGGGGCACC 6278
1351 TGCCTCTGGGATATGTGATCTGTAGTCTACAGAAAATTGGGGTCAACCGTCTATT 1410
6279 TGCCTCTGGGATATGTGATCTGTAGTCTACAGAAAATTGGGGTCAACCGTCTATT 6338
1411 ATGGGGTCTGTGTGGAAGAGACACACACTCTATTGTCATCAGATGCTTAAG 1470
6339 ATGGGGTCTGTGTGGAAGAGACACACACTCTATTGTCATCAGATGCTTAAG 6398
1471 CATATGATACAGAGTACATATATGTTTGGGCACACATGCTGTACCAACACCCA 1530
6399 CATATGATACAGAGTACATATATGTTTGGGCACACATGCTGTACCAACACCCA 6458
1531 ACCCAAGAGTATGTTGTAATGTGACAGAAAATTTTAACTGTGGAAAATGACA 1590
6459 ACCCAAGAGTATGTTGTAATGTGACAGAAAATTTTAACTGTGGAAAATGACA 6518
1591 TGTGAGAACAGATGATGAGATATATCAGTTTATGGGATCAAGCCTTAAGCCATGTG 1650
6519 TGTGAGAACAGATGATGAGATATATCAGTTTATGGGATCAAGCCTTAAGCCATGTG 6578
1651 TAAATTTAACCCACTGTGTAGTTTAAAGTGCACATGATTTGAAGATATATA 1710
6579 TAAATTTAACCCACTGTGTAGTTTAAAGTGCACATGATTTGAAGATATATA 6638
1711 CCAATAGTAGTAGGGGAGATGATATGAGAAAAGAGATATAAACTGCTTTTCA 1770
6639 CCAATAGTAGTAGGGGAGATGATATGAGAAAAGAGATATAAACTGCTTTTCA 6698
1771 ATATCAGCACAAGCATTAGAGATTAAGTGCAGAAAAGATATGATCTTTTATAACTTG 1830
6699 ATATCAGCACAAGCATTAGAGATTAAGTGCAGAAAAGATATGATCTTTTATAACTTG 6758
1831 ATATAGTACCAATAGATATATACAGCTATAGTTGATAGTTTAACTCAGTCATTA 1890
6759 ATATAGTACCAATAGATATATACAGCTATAGTTGATAGTTTAACTCAGTCATTA 6818
1891 CACAGGCTGTCCAAAGTATCTTTGAGCCAATTTCCATACATTTATTTGCGCCGGCTG 1950
6819 CACAGGCTGTCCAAAGTATCTTTGAGCCAATTTCCATACATTTATTTGCGCCGGCTG 6878
1951 GTTTTGCATTTCTAAATGTATATATAGAGCTTCAATGGAACAGACATGTCAAAATG 2010

6879 GTTTTGCATTTCTAAATGTATATATAGAGCTTCAATGGAACAGACCATGTACAAATG 6938
2011 TCAGCAGATCAATGTACACATGGAATCAGGCCAGTAGTATCACTCACTGCTTTAA 2070
6939 TCAGCAGATCAATGTACACATGGAATCAGGCCAGTAGTATCACTCACTGCTTTAA 6998
2071 ATGGCAGTCTAGCAGAGAGATGTAGTAAATTTAGATCTGCCAATTTTCACAGACAATGCTA 2130
6999 ATGGCAGTCTAGCAGAGAGATGTAGTAAATTTAGATCTGCCAATTTTCACAGACAATGCTA 7058
2131 AAACCATATATGTCAGCTGAACACATCTGTAGAAATTAATTTGACAGACCACAACA 2190
7059 AAACCATATATGTCAGCTGAACACATCTGTAGAAATTAATTTGACAGACCACAACA 7118
2191 ATACAGAAAAGATATCCGTATCCAGAGGGGACAGGAGAGCATTTGTTTCAATAGGAA 2250
7119 ATACAGAAAAGATATCCGTATCCAGAGGGGACAGGAGAGCATTTGTTTCAATAGGAA 7178
2251 AAATAGGAATATGAGACAGACACATTTGTAACTTATAGTACGAAATGCAATGCCCTT 2310
7179 AAATAGGAATATGAGACAGACACATTTGTAACTTATAGTACGAAATGCAATGCCCTT 7238
2311 TAAACAGATAGCTAGCAAAATTAAGAGAAATTTGAAATTAATTAATTAATTAATCTTTA 2370
7239 TAAACAGATAGCTAGCAAAATTAAGAGAAATTTGAAATTAATTAATTAATTAATCTTTA 7298
2371 AGCAATCTCAGAGGGGACCCAGAAATTTGTAAGCAGATTTTAATTTGGAGGGGAA 2430
7299 AGCAATCTCAGAGGGGACCCAGAAATTTGTAAGCAGATTTTAATTTGGAGGGGAA 7358
2431 TTTTCTACTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2490
7359 TTTTCTACTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7418
2491 CTGAAGGGTCAAAATTAACACTGTAAGAGAACTGACACATCACTCCCATGAGAAATTAAC 2550
7419 CTGAAGGGTCAAAATTAACACTGTAAGAGAACTGACACATCACTCCCATGAGAAATTAAC 7478
2551 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2610
7479 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7538
2611 AATTTAGATGTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2670
7539 AATTTAGATGTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7598
2671 ACAATGGGTCCGAGATCTTCAAGACCTGAGAGAGGAGATATGAGGACAAATTTGAGAAATG 2730
7599 ACAATGGGTCCGAGATCTTCAAGACCTGAGAGAGGAGATATGAGGACAAATTTGAGAAATG 7658
2731 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2790
7659 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7718
2791 AGAGAAAGTGTGTCAG 2850
7719 AGAGAAAGTGTGTCAG 7778
2851 TCTTGGGAGCAG 2910
7779 TCTTGGGAGCAG 7838
2911 GACAAATTTTGTCTGATATAGTGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2970
7839 GACAAATTTTGTCTGATATAGTGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7898
2971 AACAGCATCTGTTCGACATCTCAGAGTGTGGGACATCAAAACAGCTCCAGCAAGAAATCCGCG 3030
7899 AACAGCATCTGTTCGACATCTCAGAGTGTGGGACATCAAAACAGCTCCAGCAAGAAATCCGCG 7958
3031 CTGTGGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3090

QY 2191 ATACAGAAAAAGTATCCGTATCCAGAGGGGACCGAGAGACATTTGTACAAATGGAA 2250
 |||||||
 Db 7119 ATACAGAAAAAGTATCCGTATCCAGAGGGGACCGAGAGACATTTGTACAAATGGAA 7178
 QY 2251 AAATAGAAATATGAGACACACATTTACATTTAGAGCAAAATGGAATGCGACTT 2310
 |||||||
 Db 7179 AAATAGAAATATGAGACACATTTACATTTAGAGCAAAATGGAATGCGACTT 7238
 QY 2311 TAAACAGATAGCTAGCAAAATTTAGAGAACATTTGAAATATATAAACAATATCTTTA 2370
 |||||||
 Db 7239 TAAACAGATAGCTAGCAAAATTTAGAGAACATTTGAAATATATAAACAATATCTTTA 7298
 QY 2371 AGCAATCCGAGAGGAGGACCCAGAAATTTGTAAGCACAGTTTATTTGTGGAGGGGAA 2430
 |||||||
 Db 7299 AGCAATCCGAGAGGAGGACCCAGAAATTTGTAAGCACAGTTTATTTGTGGAGGGGAA 7358
 QY 2431 TTTTCTACTGTATTTCAACACACTGTTTAAATAGTACTGTTTAAATAGTACTTGGAGTA 2490
 |||||||
 Db 7359 TTTTCTACTGTATTTCAACACACTGTTTAAATAGTACTGTTTAAATAGTACTTGGAGTA 7418
 QY 2491 CTGAAGGTCCTAAATACACTGAGAGAGTGCACACATCCCATGCGAGAAATAAAC 2550
 |||||||
 Db 7419 CTGAAGGTCCTAAATACACTGAGAGAGTGCACACATCCCATGCGAGAAATAAAC 7478
 QY 2551 AATTTATTAACATGTGGCAGAGAGTGAAGAAAAAGCAATGTATGCCCCCATCAGTGGAC 2610
 |||||||
 Db 7479 AATTTATTAACATGTGGCAGAGAGTGAAGAAAAAGCAATGTATGCCCCCATCAGTGGAC 7538
 QY 2611 AATTTAGATGTTTCATCAATATTTACTGGGCTGCTATTAACAAGAGATGGTGTATTAACA 2670
 |||||||
 Db 7539 AATTTAGATGTTTCATCAATATTTACTGGGCTGCTATTAACAAGAGATGGTGTATTAACA 7598
 QY 2671 ACAATGGGTCGAGATCTTCAGACCTGGAGAGGCGATATGAGGACATTTGGAAGATG 2730
 |||||||
 Db 7599 ACAATGGGTCGAGATCTTCAGACCTGGAGAGGCGATATGAGGACATTTGGAAGATG 7658
 QY 2731 AATTTATTAATATTAAGTATTAAGTAAATTTGAACCAATTTGAGATAGCACCCACCAAGGCA 2790
 |||||||
 Db 7659 AATTTATTAATATTAAGTATTAAGTAAATTTGAACCAATTTGAGATAGCACCCACCAAGGCA 7718
 QY 2791 AGAGAGAGTGGTCGAGAGAGAAAAAGAGCAGTGGGAATAGAGACTTTGTTCTTGGGT 2850
 |||||||
 Db 7719 AGAGAGAGTGGTCGAGAGAGAAAAAGAGCAGTGGGAATAGAGACTTTGTTCTTGGGT 7778
 QY 2851 TCTTGGGAGCAGAGAGAGACCTATGGGCTGCACGTCAATGACGCTGACGGTACAGGCCA 2910
 |||||||
 Db 7779 TCTTGGGAGCAGAGAGAGACCTATGGGCTGCACGTCAATGACGCTGACGGTACAGGCCA 7838
 QY 2911 GACATTTATTTGCTGATATAGTACGAGACAGAAACATTTGCTGAGGGCTATTTGAGGGC 2970
 |||||||
 Db 7839 GACATTTATTTGCTGATATAGTACGAGACAGAAACATTTGCTGAGGGCTATTTGAGGGC 7898
 QY 2971 AACAGCATCTGTTGCACTACAGCTCTGGGGCATCAACAGCTCCAGGCAAGAAATCTGG 3030
 |||||||
 Db 7899 AACAGCATCTGTTGCACTACAGCTCTGGGGCATCAACAGCTCCAGGCAAGAAATCTGG 7958
 QY 3031 CTGTGGAAGATACTTAAAGATCAACAGCTCTGGGGATTTGGGCTGCTGGAAGC 3090
 |||||||
 Db 7959 CTGTGGAAGATACTTAAAGATCAACAGCTCTGGGGATTTGGGCTGCTGGAAGC 8018
 QY 3091 TCATTTGACACCTGCTGCTGGCTTGGAATGCTAGTTGGATTAATTAATCTGGAAGC 3150
 |||||||
 Db 8019 TCATTTGACACCTGCTGCTGGCTTGGAATGCTAGTTGGATTAATTAATCTGGAAGC 8078
 QY 3151 TTTGGAATTAACATGACTGATGAGTGGAGAGAGAAATTAACAAATTAACAAGCTTAA 3210
 |||||||
 Db 8079 TTTGGAATTAACATGACTGATGAGTGGAGAGAGAAATTAACAAATTAACAAGCTTAA 8138
 QY 3211 TACACTCTTAATTAAGATATGCAAAACACAGAAAGAAAGATTAACAAGATTTATGG 3270
 |||||||
 Db 8139 TACACTCTTAATTAAGATATGCAAAACACAGAAAGAAAGATTAACAAGATTTATGG 8198

QY 3271 AATTAGTAATATGGGCAAGTTTGTGAATGGTTTAACATAACAAATGGCTGTGATTA 3330
 |||||||
 Db 8199 AATTAGTAATATGGGCAAGTTTGTGAATGGTTTAACATAACAAATGGCTGTGATTA 8258
 QY 3331 TAAATTTATTAATATGATAGTAGAGGCTTGTGATTTAAGAAATAGTTTGTGCTTAC 3390
 |||||||
 Db 8259 TAAATTTATTAATATGATAGTAGAGGCTTGTGATTTAAGAAATAGTTTGTGCTTAC 8318
 QY 3391 TTTCTATATGTAATATGATAGTAGAGGCTTGTGATTTAAGAAATAGTTTGTGCTTAC 3450
 |||||||
 Db 8319 TTTCTATATGTAATATGATAGTAGAGGCTTGTGATTTAAGAAATAGTTTGTGCTTAC 8378
 QY 3451 CAATCCGAGAGGAGCCGAGACAGGCCGGAAGAAATGAAGAAGTAAGCTGAGAGAGACA 3510
 |||||||
 Db 8379 CAATCCGAGAGGAGCCGAGACAGGCCGGAAGAAATGAAGAAGTAAGCTGAGAGAGACA 8438
 QY 3511 GAGACAGATCCATTTGATTAAGTAAGAGATCTTACACTTATCTGGAAGATCTGGGA 3570
 |||||||
 Db 8439 GAGACAGATCCATTTGATTAAGTAAGAGATCTTACACTTATCTGGAAGATCTGGGA 8498
 QY 3571 GCGTGGCTCTTTCAGTACACACCGCTTGAGAGACTTACTGTTGATTAACGAGAGATTG 3630
 |||||||
 Db 8499 GCGTGGCTCTTTCAGTACACACCGCTTGAGAGACTTACTGTTGATTAACGAGAGATTG 8558
 QY 3631 TGAACCTCTGAGAGCAGAGGGGTTGGAAAGCCCTCAAAATATGCTGGAATCTCTACAGT 3690
 |||||||
 Db 8559 TGAACCTCTGAGAGCAGAGGGGTTGGAAAGCCCTCAAAATATGCTGGAATCTCTACAGT 8618
 QY 3691 ATTTGAGTCAGAACTAAAGAAATAGTCTGTTAACTTGCTCAATGCCACAGCCATAGCAG 3750
 |||||||
 Db 8619 ATTTGAGTCAGAACTAAAGAAATAGTCTGTTAACTTGCTCAATGCCACAGCCATAGCAG 8678
 QY 3751 TAGCTGAGGGGACATATAGGCTTATAGAAATTTCAAGACCTTATAGCTATTCGCC 3810
 |||||||
 Db 8679 TAGCTGAGGGGACATATAGGCTTATAGAAATTTCAAGACCTTATAGCTATTCGCC 8738
 QY 3811 ACATACCTAGAAAGATTAAGACAGAGGCTTGAAGAATTTGCTATTAAGTGGTGGCAAG 3870
 |||||||
 Db 8739 ACATACCTAGAAAGATTAAGACAGAGGCTTGAAGAATTTGCTATTAAGTGGTGGCAAG 8798
 QY 3971 CCAGCAGCAGATGGGGTGGGACGATCTGAGATCTGAGACTTGA 3971
 |||||||
 Db 8859 CCAGCAGCAGATGGGGTGGGACGATCTGAGACTTGA 8899
 RESULT 5
 ; US-08-646-538-35
 ; Sequence 35, Application US/08646538
 ; Patent No. 6027881
 ; GENERAL INFORMATION:
 ; APPLICANT: Pavlakis, George N.
 ; APPLICANT: Galtanaris, George A.
 ; APPLICANT: Stauber, Roland H.
 ; APPLICANT: Yournakis, John N.
 ; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
 ; NUMBER OF SEQUENCES: 37
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, 8th Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIA TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

Db 7959 CTGTGAAAAGATACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTCCTGGAAAAC 8018
 QY 3091 TCATTTGACACCACTCTGCTTGGAAATGCTAGTTGGATTAATTAATCTGGAACAGA 3150
 Db 8019 TCATTTGACACCACTCTGCTTGGAAATGCTAGTTGGATTAATTAATCTGGAACAGA 8078
 QY 3151 TTTGGAAATTAACATGACCTGGATGGATGGGACAGAAAATTAACAATTACACAGCTTAA 3210
 Db 8079 TTTGGAAATTAACATGACCTGGATGGATGGGACAGAAAATTAACAATTACACAGCTTAA 8138
 QY 3211 TACACTCTTAAATTAAGAGATGCAAAACAGCAAGAAAAGATGACAGATTAATGG 3270
 Db 8139 TACACTCTTAAATTAAGAGATGCAAAACAGCAAGAAAAGATGACAGATTAATGG 8198
 QY 3271 AATTAATTAAGAGATGCAAGTTGGTAAATGTTAACTAACAATGGCTGTGTATA 3330
 Db 8199 AATTAATTAAGAGATGCAAGTTGGTAAATGTTAACTAACAATGGCTGTGTATA 8258
 QY 3331 TAAATTAATTAAGATAGTAGAGAGCTTGTAGTTTAAGAAATAGTTTGGTGTAC 3390
 Db 8259 TAAATTAATTAAGATAGTAGAGAGCTTGTAGTTTAAGAAATAGTTTGGTGTAC 8318
 QY 3391 TTTCTAATGTAATAGAGATGACAGAGATTAATCAACATTAATGTTTACAGACCACTCC 3450
 Db 8319 TTTCTAATGTAATAGAGATGACAGAGATTAATCAACATTAATGTTTACAGACCACTCC 8378
 QY 3451 CAATCCGAGGGGACCCGACAGCCCGAAGAAATGAAGAAACAGGTGAGAGAGACA 3510
 Db 8379 CAATCCGAGGGGACCCGACAGCCCGAAGAAATGAAGAAACAGGTGAGAGAGACA 8438
 QY 3511 GAGACAGATCCATTCGATTAAGTAAGAGATCCTTACACTTATCTGGAGCATCTGGCA 3570
 Db 8439 GAGACAGATCCATTCGATTAAGTAAGAGATCCTTACACTTATCTGGAGCATCTGGCA 8498
 QY 3571 GCCTTGCTCTTCACTACACACCCCTTGAGAGACTTACTTGTGATTAACAGAGATTG 3630
 Db 8499 GCCTTGCTCTTCACTACACACCCCTTGAGAGACTTACTTGTGATTAACAGAGATTG 8558
 QY 3631 TGGAACTCTGGGAGGAGGGGGTGGGAAAGCCCAATATGAGAGATTCCTACAGT 3690
 Db 8559 TGGAACTCTGGGAGGAGGGGGTGGGAAAGCCCAATATGAGAGATTCCTACAGT 8618
 QY 3691 ATTGAGTCAGAGATTAAGAAATAGTGTGTTAACTGCTCAATCCACAGCATAGACG 3750
 Db 8619 ATTGAGTCAGAGATTAAGAAATAGTGTGTTAACTGCTCAATCCACAGCATAGACG 8678
 QY 3751 TAGCTAGGGGACAGATAGGGTTATGAAGATTAACAAGCAGCTTATAGAGCTATTGCC 3810
 Db 8679 TAGCTAGGGGACAGATAGGGTTATGAAGATTAACAAGCAGCTTATAGAGCTATTGCC 8738
 QY 3811 ACATACCTAAGAAATTAAGACAGGGCTTGGAAAGATTTGCTATTAAGATGGGTGGCAAG 3870
 Db 8739 ACATACCTAAGAAATTAAGACAGGGCTTGGAAAGATTTGCTATTAAGATGGGTGGCAAG 8798
 QY 3871 TGGTAAAAAAGTAGTGTGTTGATGGCTGCTGAAGGAAAGATGAGAGAGATGAG 3930
 Db 8799 TGGTAAAAAAGTAGTGTGTTGATGGCTGCTGAAGGAAAGATGAGAGAGATGAG 8858
 QY 3931 CCAGCAGCAGATGGGGTGGAGCAGTA 3957
 Db 8859 CAAGAAATGGCTAGCAAGAGAGAGAA 8885

RESULT 6
 US-09-503-222-35
 ; Sequence 35, Application US/09503222
 ; Patent No. 6265548
 ; GENERAL INFORMATION:
 ; APPLICANT: Pavlakis, George N.
 ; APPLICANT: Galanaris, George A.
 ; APPLICANT: Stauder, Roland H.
 ; APPLICANT: Vournakis, John N.
 ; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent

; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
 ; NUMBER OF SEQUENCES: 37
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, 8th floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/503,222
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/646,538
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Weber, Kenneth A.
 ; REGISTRATION NUMBER: 31,677
 ; REFERENCE/DOCKET NUMBER: 015280-249000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 35:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 15581 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA
 ; FEATURE:
 ; NAME/KEY:
 ; LOCATION: 1..15581
 ; OTHER INFORMATION: /note="pNlnSG11"
 ; US-09-503-222-35

Query Match 42.3%; Score 2636.6; DB 3; Length 15581;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 2648; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1291 CCATGAGCTGAAGAGAAATGATACAGCACTTGTGAGATGGGGTGAATGGGCACCA 1350
 Db 6219 CAATGAGCTGAAGAGAAATGATACAGCACTTGTGAGATGGGGTGAATGGGCACCA 6278
 QY 1351 TGCCTCGGGATATTGATGATCTGATGCTACAGAAAATTTGGGGTCCACGCTATT 1410
 Db 6279 TGCCTCGGGATATTGATGATCTGATGCTACAGAAAATTTGGGGTCCACGCTATT 6338
 QY 1411 ATGGGGTACCTGTGGAAGGAAGCAACCACTCTATTGTCATCAGATGCTAAG 1470
 Db 6339 ATGGGGTACCTGTGGAAGGAAGCAACCACTCTATTGTCATCAGATGCTAAG 6398
 QY 1471 CATATGATCAGAGGTACATTAATGTTGGGCCACACATGCTGTATCCACAGACCCA 1530
 Db 6399 CATATGATCAGAGGTACATTAATGTTGGGCCACACATGCTGTATCCACAGACCCA 6458
 QY 1531 ACCCAAGAAATAGTATTGTTGAATGTGACAGAAAATTTTAACATGTTGGAATAATGACA 1590
 Db 6459 ACCCAAGAAATAGTATTGTTGAATGTGACAGAAAATTTTAACATGTTGGAATAATGACA 6518
 QY 1591 TGGTAGAACAGATGATGAGATATATACGTTTATGGATCAAAAGCCATAAGCCATGTG 1650
 Db 6519 TGGTAGAACAGATGATGAGATATATACGTTTATGGATCAAAAGCCATAAGCCATGTG 6578
 QY 1651 TAAATTAACCCCACTCTGTGTAGTTTAAAGTGCACATGATTTGAAGATGATTAATA 1710
 Db 6579 TAAATTAACCCCACTCTGTGTAGTTTAAAGTGCACATGATTTGAAGATGATTAATA 6638

QY 1711 CCAATAGTAGAGCGGGAGAAATGATATATGAGAAAAGAGATATAAAAAAAGTCTCTTTCA 1770
|||||
Db CCAATAGTAGAGCGGGAGAAATGATATGAGAAAAGAGATATAAAAAAAGTCTCTTTCA 6698
QY 1771 ATATCAGACAAGAGATAGAGATATAGGTGCGAGAAAAGAAATATGCATCTTTTATAACTTG 1830
|||||
Db ATATCAGACAAGAGATAGAGATATAGGTGCGAGAAAAGAAATATGCATCTTTTATAACTTG 6758
QY 1831 ATATAGTACCAATAGATATATACCGCTATAGTGTGATATGTATACCTCAGTCATTA 1890
|||||
Db ATATAGTACCAATAGATATATACCGCTATAGTGTGATATGTATACCTCAGTCATTA 6818
QY 1891 CACAGGCTGTCCAAAGGTATCTTTGAGCCAAATTCCTACATTTATTTGTGCGCGGTG 1950
|||||
Db CACAGGCTGTCCAAAGGTATCTTTGAGCCAAATTCCTACATTTATTTGTGCGCGGTG 6878
QY 1951 GTTTTGCATCTTAAATGATATATATAGCTTCAATGAAACGAGACCATGTACAAATG 2010
|||||
Db GTTTTGCATCTTAAATGATATATATAGAGCTCAATGAAACGAGACCATGTACAAATG 6938
QY 2011 TCAGCAGACGATCAATGTACACATGGGAATCAGGCCAGTATCTCAACTCAGCTGTAA 2070
|||||
Db TCAGCAGACGATCAATGTACACATGGGAATCAGGCCAGTATCTCAACTCAGCTGTAA 6998
QY 2071 ATGCGAGTCTAGCAGAGAGAGATGTATTAATTAGATCTGCCAATTTTCACAGCAATGCTA 2130
|||||
Db ATGCGAGTCTAGCAGAGAGAGATGTATTAATTAGATCTGCCAATTTTCACAGCAATGCTA 7058
QY 2131 AACCATATATATAGTACAGTGAACACATCTGTATGAAATTAATTTGTCAAGACCCAAACA 2190
|||||
Db AACCATATATATAGTACAGTGAACACATCTGTATGAAATTAATTTGTCAAGACCCAAACA 7118
QY 2191 ATACAGAAAAAGTATCCGTATCCAGAGGGAGCCAGGAGAGCATTTGTTCAATATAGAA 2250
|||||
Db ATACAGAAAAAGTATCCGTATCCAGAGGGAGCCAGGAGAGCATTTGTTCAATATAGAA 7178
QY 2251 AAAATAGGAATATAGAGACAACCATTTTACATTAATAGTAGCAAAAAAGGAATGCGACTT 2310
|||||
Db AAAATAGGAATATAGAGACAACCATTTTACATTAATAGTAGCAAAAAAGGAATGCGACTT 7238
QY 2311 TAAACAGATAGTACGAAATTAAGAGAACATTTTGAATATATTAACATTAATCTTTA 2370
|||||
Db TAAACAGATAGTACGAAATTAAGAGAACATTTTGAATATATTAACATTAATCTTTA 7298
QY 2371 AGCAATCTCAGAGAGGGAGCCAGAAATTTGAACGACACAGTTTAATTTGTGAGGGAGAT 2430
|||||
Db AGCAATCTCAGAGAGGGAGCCAGAAATTTGAACGACACAGTTTAATTTGTGAGGGAGAT 7358
QY 2431 TTTTCTACTGTATTTCAACACAACTGTTAATAGTACTGGTTAATAGTACTTGAGATA 2490
|||||
Db TTTTCTACTGTATTTCAACACAACTGTTAATAGTACTGGTTAATAGTACTTGAGATA 7418
QY 2491 CTGAAGGGTCAAAATPAACCTGAGAGAGTGAACACATCCCATGCGAATTAAC 2550
|||||
Db CTGAAGGGTCAAAATPAACCTGAGAGAGTGAACACATCCCATGCGAATTAAC 7478
QY 2551 AATTTATTAACATGTGCGAGAGAGTAGGAAAAAGCATGTATGCCCTCCCATCACTGAGAC 2610
|||||
Db AATTTATTAACATGTGCGAGAGAGTAGGAAAAAGCATGTATGCCCTCCCATCACTGAGAC 7538
QY 2611 AAATTAGATGTTCAATCAATATTTACTGGGTGCTATTAACAAGAGATGGTGTAAATACA 2670
|||||
Db AAATTAGATGTTCAATCAATATTTACTGGGTGCTATTAACAAGAGATGGTGTAAATACA 7598
QY 2671 ACAATGGGTGAGATCTTACAGCTGAGAGAGGCGATATAGAGGCAATTTGAGAAATG 2730
|||||
Db ACAATGGGTGAGATCTTACAGCTGAGAGAGGCGATATAGAGGCAATTTGAGAAATG 7658
QY 2731 AATTAATATTAATTAAGATAGTAAAAATTTGAACATTTGAGATAGACCCCAAGAGCAA 2790
|||||
Db AATTAATATTAATTAAGATAGTAAAAATTTGAACATTTGAGATAGACCCCAAGAGCAA 7718
QY 2791 AGAGAAAGATGGTGCAGAGAGAAAAAGAGCATGTGGGAATAGAGCTTTGTTCTTGGGT 2850
|||||

Db 7719 AGAGAAAGATGGTGCAGAGAGAAAAAAGACAGATGGGAATAGAGCTTTGTTCTTGGGT 7778
|||||
QY 2851 TCTTGGGAGCAGAGAGAAAGCACTATGGCTGCAGTCAATATACGCTGACGGGTACAGGCA 2910
|||||
Db TCTTGGGAGCAGAGAGAAAGCACTATGGCTGCAGTCAATATACGCTGACGGGTACAGGCA 7838
QY 2911 GACAATTAATGCTGATATAGTGAAGCAGCAGAAACAAATTTCTGAGGCTATTTAGGCGC 2970
|||||
Db GACAATTAATGCTGATATAGTGAAGCAGCAGAAACAAATTTCTGAGGCTATTTAGGCGC 7898
QY 2971 AACAGCATCTGTGCACTCAACAGTCTGGGCAATCAACAGCTCCAGGCAAGAAATCTGCG 3030
|||||
Db AACAGCATCTGTGCACTCAACAGTCTGGGCAATCAACAGCTCCAGGCAAGAAATCTGCG 7958
QY 3031 CTGTGAGAGATPACTTAAGATCAACAGCTCCGAGGATTTGGGGTGTCTGTGAANAC 3090
|||||
Db CTGTGAGAGATPACTTAAGATCAACAGCTCCGAGGATTTGGGGTGTCTGTGAANAC 8018
QY 3091 TCATTTGCACCACTGCTGTGCTTGGAAATGCTAGTTGAGATTAATTAATCTGTGAAACA 3150
|||||
Db TCATTTGCACCACTGCTGTGCTTGGAAATGCTAGTTGAGATTAATTAATCTGTGAAACA 8078
QY 3151 TTTGGAAATACATGACCTGATGGAAGTGGAGACAGAAATTAACAAATTTACAAAGCTTAA 3210
|||||
Db TTTGGAAATACATGACCTGATGGAAGTGGAGACAGAAATTAACAAATTTACAAAGCTTAA 8138
QY 3211 TACACTCTTAAATTAAGAAATCGCAAAACGAGAAAGAAAGAAACAAATTAATTTGG 3270
|||||
Db TACACTCTTAAATTAAGAAATCGCAAAACGAGAAAGAAAGAAACAAATTAATTTGG 8198
QY 3271 AATTAGATTAATGGGCAAGTTTGTGAATTTGTTTAAATTAACAAATTTGGCTGTGTATA 3330
|||||
Db AATTAGATTAATGGGCAAGTTTGTGAATTTGTTTAAATTAACAAATTTGGCTGTGTATA 8258
QY 3331 TAAATTAATATATATAGTATAGAGGCTGTGATAGGTTTAAGATTAATTTTCTGTAC 3390
|||||
Db TAAATTAATATATATAGTATAGAGGCTGTGATAGGTTTAAGATTAATTTTCTGTAC 8318
QY 3391 TTTCTATGATTAATAGATTAAGTATAGGAGGGATTTTACATTTATCTTTAGACCCACCTCC 3450
|||||
Db TTTCTATGATTAATAGATTAAGTATAGGAGGGATTTTACATTTATCTTTAGACCCACCTCC 8378
QY 3451 CAATCCGAGGGGAGCCGACAGGCCCGGAAGAAATAGAAAGAAAGGTGAGAGAGACA 3510
|||||
Db CAATCCGAGGGGAGCCGACAGGCCCGGAAGAAATAGAAAGAAAGGTGAGAGAGACA 8438
QY 3511 GAGACAGATCCATTCGATTAAGTGAACGGATCTCTTAGACCTATCTGGGACGATCTGCGGA 3570
|||||
Db GAGACAGATCCATTCGATTAAGTGAACGGATCTCTTAGACCTATCTGGGACGATCTGCGGA 8498
QY 3571 GCCTGTGCTCTTACAGTACACCGCTTGAGAGACTTACTCTTGATTTTAACGAGAGATG 3630
|||||
Db GCCTGTGCTCTTACAGTACACCGCTTGAGAGACTTACTCTTGATTTTAACGAGAGATG 8558
QY 3631 TGAACCTTCTGGAGCGCAGGGGGTGGGAAGCCCTCAATATTTGTGTGAATCTCTACAGT 3690
|||||
Db TGAACCTTCTGGAGCGCAGGGGGTGGGAAGCCCTCAATATTTGTGTGAATCTCTACAGT 8618
QY 3691 ATTGGAGTCAGAGAACTAAGAAATAGTGTAACTTGCTCAATGCCACAGCATATAGCAG 3750
|||||
Db ATTGGAGTCAGAGAACTAAGAAATAGTGTGTAACTTGCTCAATGCCACAGCATATAGCAG 8678
QY 3751 TAGCTGAGGGAGCAGATATGGGTTATAGAAATTAACAAGCAGCTTATAGACTATTTGCC 3810
|||||
Db TAGCTGAGGGAGCAGATATGGGTTATAGAAATTAACAAGCAGCTTATAGACTATTTGCC 8738
QY 3811 ACATACCTAGAAAGATTAAGACAGGGCTTGAAAGGATTTTCTATAGATGGGTGCGAAG 3870
|||||
Db ACATACCTAGAAAGATTAAGACAGGGCTTGAAAGGATTTTCTATAGATGGGTGCGAAG 8798
QY 3871 TGCTCAAAAAGTACTGTGATTTGATGAGCTGCTGTGAAGGAAAGATGACGAGCTGAG 3930
|||||

|||||
Db 7138 TTGGGTTCTTGGGAGCAGCAGAGAACTATGGCGCGCAGCGTCAATGACGCTGACGCTAC 7197
OY 2905 AGGCGAGCAATTTATGCTGATATAGTGCAGCAGCAGCAACAATTTGCTGAGGCTATTG 2964
Db 7198 AGGCGAGCAATTTATGCTGATATAGTGCAGCAGCAGCAACAATTTGCTGAGGCTATTG 7257
OY 2965 AGGCGAGCAATTTATGCTGATATAGTGCAGCAGCAGCAACAATTTGCTGAGGCTATTG 3024
Db 7258 AGGCGAGCAATTTATGCTGATATAGTGCAGCAGCAGCAACAATTTGCTGAGGCTATTG 7317
OY 3025 TCCTGGCTGTGAGAAAGATACCTAAAGATCAACACCTCTGGGATTTGGGGTTCCTG 3084
Db 7318 TCCTGGCTGTGAGAAAGATACCTAAAGATCAACACCTCTGGGATTTGGGGTTCCTG 7377
OY 3085 GAAACATCTTGCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3144
Db 7378 GAAACATCTTGCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7437
OY 3145 AACAGATTGGAAATACATGACCTGATGAGTGGGAGAGAAATTAACAATTACAGAA 3204
Db 7438 AACAGATTGGAAATACATGACCTGATGAGTGGGAGAGAAATTAACAATTACAGAA 7497
OY 3205 GCTTAATACACTCTCTTAATTAAGAAATCCAAACCAAGAAAGAAAGAAAGAAAGAAAT 3264
Db 7498 GCTTAATACACTCTCTTAATTAAGAAATCCAAACCAAGAAAGAAAGAAAGAAAGAAAT 7557
OY 3265 TATTGAATTTAGTAAATGGGCAAGTTGGTGAATTTGGTAACTAACTAACTAACTGCTG 3324
Db 7558 TATTGAATTTAGTAAATGGGCAAGTTGGTGAATTTGGTAACTAACTAACTAACTGCTG 7617
OY 3325 GGTATATATAATTTATTCATTAATGATAGTAGAGGCTTGGTAACTAACTAACTGCTG 3384
Db 7618 GGTATATATAATTTATTCATTAATGATAGTAGAGGCTTGGTAACTAACTAACTGCTG 7677
OY 3385 CTGACTTTCTAATGATAGTAAATGAGTAGGCAAGGATATTCACCATTTATGCTTACAGCC 3444
Db 7678 CTGACTTTCTAATGATAGTAAATGAGTAGGCAAGGATATTCACCATTTATGCTTACAGCC 7737
OY 3445 ACCCTCCCAATCCCGAGGGGAGCCGACAGGCGCGGAGAAATTAAGAAAGAAAGTGGAGGA 3504
Db 7738 ACCCTCCCAATCCCGAGGGGAGCCGACAGGCGCGGAGAAATTAAGAAAGAAAGTGGAGGA 7797
OY 3505 GAGACAGAGACAGATCCATTCATTAATGATAGTAGGCAAGGATATTCACCATTTATGCTTACAG 3564
Db 7798 GAGACAGAGACAGATCCATTCATTAATGATAGTAGGCAAGGATATTCACCATTTATGCTTACAG 7857
OY 3565 TCGGAGGCTGCTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 3624
Db 7858 TCGGAGGCTGCTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 7917
OY 3625 GGATTGTTGGAATCTTGGGAGCGCAGGCGGAGGAGAAAGCTTAAATTAATTTGGGAAATCC 3684
Db 7918 GGATTGTTGGAATCTTGGGAGCGCAGGCGGAGGAGAAAGCTTAAATTAATTTGGGAAATCC 7977
OY 3685 TACAGTATTTGAGTACAGAACTAAGATAGTCTGTTAACTTGTCTCAATCCACAGCA 3744
Db 7978 TACAGTATTTGAGTACAGAACTAAGATAGTCTGTTAACTTGTCTCAATCCACAGCA 8037
OY 3745 TAGCAGTATTTGAGTACAGAACTAAGATAGTCTGTTAACTTGTCTCAATCCACAGCA 3804
Db 8038 TAGCAGTATTTGAGTACAGAACTAAGATAGTCTGTTAACTTGTCTCAATCCACAGCA 8097
OY 3805 TTGCGCACATTCCTAGAAAGTAAAGATAGAGGCTTGAAGAGTAAATTTGCTATTAATAGGCT 3864
Db 8098 TTGCGCACATTCCTAGAAAGTAAAGATAGAGGCTTGAAGAGTAAATTTGCTATTAATAGGCT 8157
OY 3865 GGCAGTATTTGAGTACAGAACTAAGATAGTCTGTTAACTTGTCTCAATCCACAGCA 3924
Db 8158 GGCAGTATTTGAGTACAGAACTAAGATAGTCTGTTAACTTGTCTCAATCCACAGCA 8217
OY 3925 GCTGAGCCAGCAGATGGGGTGGAGCAGTATCTGAGATCTAGA 3971
|||||

Db 8218 GCTGAGCCAGCAGCAGATGGGGTGGAGCAGCATCTCGAGACTAGA 8264
RESULT 8
US-08-463-210-4
Sequence 4, Application US/08463210
Patent No. 6001977
GENERAL INFORMATION:
APPLICANT: CHANG, Nancy T.
APPLICANT: GALLO, Robert C.
APPLICANT: WONG-STALL, Flossie
TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463, 210
FILING DATE: 05-JUN-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/693, 866
FILING DATE: 23-JAN-1985
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/659, 339
FILING DATE: 10-OCT-1984
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4193052
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 8933 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: HTLV-III
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..8933
OTHER INFORMATION: /standard.name="Clone BH10"
OTHER INFORMATION: /note="Corresponds to nucleotide positions 222 to
OTHER INFORMATION: 9154 in figure 3 of EP 85307260"
FEATURE:
NAME/KEY: mat.peptide
LOCATION: 113..1648
OTHER INFORMATION: /product="gag"
FEATURE:
NAME/KEY: mat.peptide
LOCATION: 1408..4452
OTHER INFORMATION: /product="pol"
FEATURE:
NAME/KEY: mat.peptide
LOCATION: 4367..4975
OTHER INFORMATION: /product="sor"
FEATURE:
NAME/KEY: mat.peptide
LOCATION: 5560..8148

QY	3365	CTGTACTTTCTATAGTGAATAGAGTTAGGACAGGGATATTCCACCAATTATCGTTTCAGACC	3444
Dd	7679	CTGTACTTTCTGTACTGTAATAGAGTTAGGACAGGGATATTCCACCAATTATCGTTTCAGACC	7738
QY	3445	ACCTCCCATCCTCCGAGGGGACCCGCACAGGCCCGAAGAATAATGAAGAACAGGTGAGAGA	3504
Dd	7739	ACCTCCCAATCCCGAGGGGACCCGCACAGGCCCGAAGAATAATGAAGAACAGGTGAGAGA	7798
QY	3505	GAGACAGAGACAGATCCATTGCGATTAGTGAAACGATCCTTAGCACTTATCTGGACGATC	3564
Dd	7799	GAGACAGAGACAGATCCATTGCGATTAGTGAAACGATCCTTAGCACTTATCTGGACGATC	7858
QY	3565	TGCGAGAGCTTGCTCTTCACAGCTACCAACCCCTTGAGAGACTTACTCTTGATTGTACGA	3624
Dd	7859	TGCGAGAGCTTGCTCTTCACAGCTACCAACCCCTTGAGAGACTTACTCTTGATTGTACGA	7918
QY	3625	GGATTGTGAACCTTCTGGAGACGAGGGGGTGGGAAAGCCCTCAATATATGATGGAATCCC	3684
Dd	7919	GGATTGTGAACCTTCTGGAGACGAGGGGGTGGGAAAGCCCTCAATATATGATGGAATCCC	7978
QY	3685	TACAGTATTTGAGTCAGGAGACTTAAAGAAATAGTGTCTTTAACTTGGTCCATGCCACAGCA	3744
Dd	7979	TACAGTATTTGAGTCAGGAGACTTAAAGAAATAGTGTCTTTAACTTGGTCCATGCCACAGCA	8038
QY	3745	TAGCAGTAGCTGAGGGGACACATAGGGTTATAGAAGTTTACAAGACCTTATAGAGCTA	3804
Dd	8039	TAGCAGTAGCTGAGGGGACACATAGGGTTATAGAAGTTTACAAGACCTTATAGAGCTA	8098
QY	3805	TYCGCACATACCTCTGAAGAATAAAGACACAGGCTTGGAAGAAGATTTCGTATAAGATGGGT	3864
Dd	8099	TYCGCACATACCTCTGAAGAATAAAGACACAGGCTTGGAAGAAGATTTCGTATAAGATGGGT	8158
QY	3865	GGCAAGTGTCTAAAAAGTAGTGTGATTGGATGGCTGCTGTAAAGGAAAATGACAGA	3924
Dd	8159	GGCAAGTGTCTAAAAAGTAGTGTGATTGGATGGCTGCTGTAAAGGAAAATGACAGA	8218
QY	3925	GCTGAGCCAGCAGCATGGGGGTGGGACCATATCTCGAGATCTTGA	3971
Dd	8219	GCTGAGCCAGCAGCATGGGGGTGGGACCATATCTCGAGATCTTGA	8265
RESULT 9			
US-09-620-958A-3			
; Sequence 3, Application US/09620958A			
; Patent No. 6294338			
; GENERAL INFORMATION:			
; APPLICANT: Nunomura, Kiyotada			
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD			
; FILE REFERENCE: GP104-02, JT			
; CURRENT APPLICATION NUMBER: US/09/620,958A			
; CURRENT FILING DATE: 2000-07-21			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 3			
; LENGTH: 8933			
; TYPE: RNA			
; ORGANISM: Human Immunodeficiency Virus			
; FEATURE:			
; NAME/KEY: source			
; LOCATION: (1)...(8933)			
; OTHER INFORMATION: Sequence of transcripts produced from the BH10			
; Patent No. 6294338			
; OTHER INFORMATION: plasmid.			
US-09-620-958A-3			
Query Match			
Best Local Similarity 41.7%; Score 2597.8; DB 3; Length 8933;			
Matches 1997; Conservative 642; Mismatches 42; Indels 6; Gaps 1;			
QY	1291	CCAATGAGAGTAAAGAGCAATATAGACACTTGTGGAGATGGGGGTGGAATGGGGCACCA	1350
Dd	5579	CAATGGAGGUGUAAAGCAAAAUAVUACGACACUUGUGGAGUUGGGUGUGAGUUGGGGCACCA	5638

QY	1331	1GCTCTTGGGATATTTGATGATCTGTACTGCTACAGAAAAATTTGGGTCAACCTCTATT	1410
Db	5639	UUCUCUCUGGGAGUUUGUAGUCUGUACUGUACAGAAAAUUUUUGGGUUCACACUADUU	5698
QY	1411	ATGGGGTACCTGTGTGGGAAGAGAACACACACACTCTTTTGTGCATCAGATGCTTAAAG	1470
Db	5659	AUGGGGGUACCTGUGUUGGAAGAGACACACACACUCUUAUUUUGGCAUCAGAGUCUAAAG	5758
QY	1471	CATATGATACAGAGGTACATATATGTTTGGGGCCACACATGCTGTACTGCCACAGACCCA	1530
Db	5759	CAUAGUACAGAGAGUACAUAAUGUUUGGGCCACACAGCCUGUGUACCCACAGACCCCA	5818
QY	1531	ACCCACGAAGATGTATTTGGTAAATGTGCACAGAAAAATTTTAACATGTGGAAAAATGACA	1590
Db	5819	ACCCCAAGAAAGUADUUUGUUUUUUGAGACAGAAAAUUUUAACUUGGGGAAAAUAGACA	5878
QY	1591	TGTTAGACACATGCGATGAGATGTATTAATCAGTTTATGTGGATCAAAAGCCTTAAGCCATGTG	1650
Db	5879	UGUUGAAGACACAUUGCAUUGAGGAAUUAACUUUUAUGGACUCAAAGCUAAAGCCANUGG	5938
QY	1651	TAAATTAACCCCACTCTGTGTATTGTTTAAAGTGCACGTATTGTAAGATGATCTATAA	1710
Db	5939	UAAAAUUUAAACCCACUCUCUGUU	5998
QY	1711	CCAATAGTAGTAGGGGGGAATGTAAATGAGAAAAAGGACATTAATAAAACCTGCTTTACA	1770
Db	5999	CCAAUAGUAGAGCGGGAGAAUAGUAAUUGAGAAAGAGACAUAAAAACUGCCUUCU	6058
QY	1771	ATATCAGCACAGCATTAAGATTAAGTGTGCGAAAAAATATGCAATCTTTTATTAACCTTG	1830
Db	6059	AUATCAGCACAGCATTAAGATTAAGTGTGCGAAAAAATATGCAATCTTTTATTAACCTTG	6118
QY	1831	ATATAGTACCAATAGATTA-----TACACACTCTTAGGTTGATTAAGTTTGTAAACCTCAG	1884
Db	6119	AUADUADUCCAUADUADUADUAGUACUACUACACUADUACUADUACUADUADUADUADUADUADU	6178
QY	1885	TCATTTACACAGGCTGTCCAAAGGTATCTCTTTGAGCCAAATTCCTCATTTATTTGTGCC	1944
Db	6179	UCAUUACACAGGCTGTCCAAAGGTATCTCTTTGAGCCAAATTCCTCATTTATTTGTGCC	6238
QY	1945	CGGCTGCTTTGGCATCTCTAAATGTATATATAAGACGTTCAATGGAACAGGACCAATGTA	2004
Db	6239	CGGCTGCTTTGGCATCTCTAAATGTATATATAAGACGTTCAATGGAACAGGACCAATGTA	6298
QY	2005	CAATGTGTACAGACATGTAACATGTACATGTAATCAGGCGCATGTATCAACTCACTGTC	2064
Db	6299	CAATGTGTACAGACATGTAACATGTACATGTAATCAGGCGCATGTATCAACTCACTGTC	6358
QY	2065	TGTTTAATGCGATCTACAGACAGAGATGTAGTAAATTAGTCTGCCAATTTCCAGACA	2124
Db	6359	TGTTTAATGCGATCTACAGACAGAGATGTAGTAAATTAGTCTGCCAATTTCCAGACA	6418
QY	2125	ATGCTAAAAACCATTAATGTACAGGTGAAACATCTGTTAAATTTAATTTGTTCAAGACCCA	2184
Db	6419	AUGGCAAAAAACCATTAATGTACAGGTGAAACATCTGTTAAATTTAATTTGTTCAAGACCCA	6478
QY	2185	ACAACAATATACAAAAAGTATCCGCTATCCAGAGGGGACCAAGGAGAGACATTTGTTCAAA	2244
Db	6479	ACAACAATATACAAAAAGTATCCGCTATCCAGAGGGGACCAAGGAGAGACATTTGTTCAAA	6538
QY	2245	TAGAAAAAATAGGAAATATGTAGACACACATTTGTAACTTATAGTAGACAAAAATGGAATG	2304
Db	6539	TAGAAAAAATAGGAAATATGTAGACACACATTTGTAACTTATAGTAGACAAAAATGGAATG	6598
QY	2305	CCACTTTAAAAACATATAGCTAGCAAAATTTAAGAGAAACAATTTGGGAAAAATTAACATTA	2364
Db	6599	ACACUUAUAAACACAUAGAUAGCAAAUUAAGAGAACAAUUAUGGAAAAUAAUAAAAACAUA	6658
QY	2365	TCTTTTAAGCAATCTCTCAGAGGGGACCCAGAAATTTGTAAAGCACATGTTTATTTGTGGAG	2424
Db	6659	UCUUUAAAGCAGUUCUUCAGAGGGGACCCAGAAAUUGUAAAGCACAGUUUUUUUUUUUUUUUU	6718

[illegible]

[illegible]

QY	3805	TTGCGCCATACCTAGAAAGTAATGACAGGGCTTGAAAAGGATTTGCTATAGATGGCT	3864
Db	8099	UUCCGCCAAVUCUAGAAAGAAVUAAGACAGGGCTUGGAAAGAAUUUUUCUAAVAGAVUGGU	8158
QY	3865	GGCAAGTGTCCAAAAGAGTGTATGTAGATGGGCTCCTGTGAACGGAAAGAAATGAGACGA	3924
Db	8159	GGCAAGUGUGUAAAAGAAAGAGUGUGUGGAGAGGCGCUCUGUAGAGGAAAGAAUGAGACGA	8218
QY	3925	GCTGAGCAGCAGCAGATGGGGGTGGGAGACAGTATCTCGAGATCTAGA	3971
Db	8219	GTUAGGCCAGCAGAGANUGGGGUGGAGACACCAUCUGAGACCUAGA	8265
RESULT 12			
US-09-700-304-1			
; Sequence 1, Application US/09700304			
; Patent No. 6500623			
; GENERAL INFORMATION:			
; APPLICANT: Genecure, LLC			
; APPLICANT: Tung, Frank Y.T.			
; TITLE OF INVENTION: Replication Defective HIV Vaccine			
; FILE REFERENCE: 00714PCTUS			
; CURRENT APPLICATION NUMBER: US/09/700.304			
; CURRENT FILING DATE: 2000-11-09			
; PRIOR APPLICATION NUMBER: PCT/US99/10523			
; PRIOR FILING DATE: 1999-05-12			
; PRIOR APPLICATION NUMBER: 60/085,115			
; PRIOR FILING DATE: 1998-05-12			
; NUMBER OF SEQ ID NOS: 1			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 1			
; LENGTH: 9719			
; TYPE: DNA			
; ORGANISM: Human Immunodeficiency virus type 1			
US-09-700-304-1			
Query Match			
Best Local Similarity 41.6%; Score 2593; DB 4; Length 9719;			
Matches 2636; Conservative 0; Mismatches 45; Indels 6; Gaps 1;			
QY	1291	CCATGAGAGTAGAGAGAGAGATATCAGCAGCTTGTGAGATGGGGGTGGAATGGGGCACCA	1350
Db	6223	CAATGAGAGTAGAAGAGAAATATCAGCAGCTTGTGAGATGGGGGTGAGATGGGGCACCA	6282
QY	1331	TGCTCCTTGGGATATGATGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1410
Db	6283	TGCTCCTTGGGATGTGTGATGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6342
QY	1411	ATGGGGTACCTGTGTGGAGAGAACCAACCACTATTTTGTGCATTCGATGCTTAAG	1470
Db	6343	ATGGGGTACCTGTGTGGAGAGAACCAACCACTATTTTGTGCATTCGATGCTTAAG	6402
QY	1471	CATATGATACAGAGGTACATTAATGTTTGGGGCACACATGCTGTGTACCCACAGACCCA	1530
Db	6403	CATATGATACAGAGGTACATTAATGTTTGGGGCACACATGCTGTGTACCCACAGACCCA	6462
QY	1531	ACCCACAGAAGTGTATTTGTTTAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1590
Db	6463	ACCCACAGAAGTGTATTTGTTTAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6522
QY	1591	TGTGTGAAGACATGATGAGATGATTAATGATTTATGAGGATCAAGGCTTAAGGCATGTG	1650
Db	6523	TGTGTGAAGACATGATGAGATGATTAATGATTTATGAGGATCAAGGCTTAAGGCATGTG	6582
QY	1651	TAAATTTAACCCACTGT	1710
Db	6583	TAAATTTAACCCACTGT	6642
QY	1711	CCAATAGTAGTAGCGGAGATGATTAATGAGAAAGAGAGATTAATAAAGTCTCTTTCA	1770
Db	6643	CCAATAGTAGTAGCGGAGATGATTAATGAGAAAGAGAGATTAATAAAGTCTCTTTCA	6702

1771 ATATCAGCAGCAAGCATAGATAGTGTGACAGAAAGATATGATCTTTTATAACTTG 1830
1772 ATATCAGCAGCAAGCATAGATAGTGTGACAGAAAGATATGATCTTTTATAACTTG 1830
6703 ATATCAGCAGCAAGCATAGATAGTGTGACAGAAAGATATGATCTTTTATAACTTG 6762
1831 ATATCAGCAGCAAGCATAGATAGTGTGACAGAAAGATATGATCTTTTATAACTTG 1884
6763 ATATCAGCAGCAAGCATAGATAGTGTGACAGAAAGATATGATCTTTTATAACTTG 6822
1885 TCATTACAGCAGCCTGTCCAAAGATATCTTTGAGCCAAATTCATACATTATGTGCCC 1944
6823 TCATTACAGCAGCCTGTCCAAAGATATCTTTGAGCCAAATTCATACATTATGTGCCC 6882
1945 CGGCTGTTTGGCATCTTAAATATGATATTAAGCTTGAACCTGAGACAGACATGTA 2004
6883 CGGCTGTTTGGCATCTTAAATATGATATTAAGCTTGAACCTGAGACAGACATGTA 6942
2005 CAAATGTGACAGCAGTACATGTACACATGTGATGAGCCAGTATGATCAACTGAC 2064
6943 CAAATGTGACAGCAGTACATGTACACATGTGATGAGCCAGTATGATCAACTGAC 7002
2065 TGTATATGAGCAGTGTACAGAGAGAGATGTATATATGATGTGCAATTTACAGACA 2124
7003 TGTATATGAGCAGTGTACAGAGAGAGATGTATATATGATGTGCAATTTACAGACA 7062
2125 ATGCTAAACCATATATAGTACAGTGAACACATCTGTGAATATTAATGTGACAGACCA 2184
7063 ATGCTAAACCATATATAGTACAGTGAACACATCTGTGAATATTAATGTGACAGACCA 7122
2185 ACAACATACAGAGAGAGAGATGCGTATCCAGAGGGAGCAGGAGACATTTGTACAA 2244
7123 ACAACATACAGAGAGAGATGCGTATCCAGAGGGAGCAGGAGACATTTGTACAA 7182
2245 TAGGAGAGAGAGAGAGATGTAGACAGACATTTGATACATTAGTACAGAGAGAGATG 2304
7183 TAGGAGAGAGAGAGAGATGTAGACAGACATTTGATACATTAGTACAGAGAGAGATG 7242
2305 CCACCTTAAACAGATAGTACAGTGAACATTTGAAGAGATTTGAAGATATTAACATTA 2364
7243 ACACCTTAAACAGATAGTACAGTGAACATTTGAAGAGATTTGAAGATATTAACATTA 7302
2365 TCTTTAAGCAATCTCAGAGAGAGAGAGAGAGAGATTTGAAGAGATTTGAAGAGATTT 2424
7303 TCTTTAAGCAATCTCAGAGAGAGAGAGAGAGAGATTTGAAGAGATTTGAAGAGATTT 7362
2425 GGGATTTTCTACTGATATTCACACACATGTTAATAGTACTGTTAATAGTACTT 2484
7363 GGGATTTTCTACTGATATTCACACACATGTTAATAGTACTGTTAATAGTACTT 7422
2485 GGAGTACTGAGAGAGAGATTAACACTGAAGAGAGATGAACATCCATCCCATGACAGAA 2544
7423 GGAGTACTGAGAGAGAGATTAACACTGAAGAGAGATGAACATCCATCCCATGACAGAA 7482
2545 TAAACCAATTTAATAACATGTGGCAGAGAGATGAAGAGAGATGTATCCCTCCCATCA 2604
7483 TAAACCAATTTAATAACATGTGGCAGAGAGATGAAGAGAGATGTATCCCTCCCATCA 7542
2605 GTGAGCAATTTAGATGTTCAATCAATATTACTGGGCTCTATTAACAGAGATGTGTTA 2664
7543 GTGAGCAATTTAGATGTTCAATCAATATTACTGGGCTCTATTAACAGAGATGTGTTA 7602
2665 ATACCAATTTAGATGTTCAATCAATATTACTGGGCTCTATTAACAGAGATGTGTTA 2724
7603 ATAGCAATTTAGATGTTCAATCAATATTACTGGGCTCTATTAACAGAGATGTGTTA 7662
2725 GAACTGAATTTATTAATATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 2784
7663 GAACTGAATTTATTAATATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 7722
2785 AGGCAAGAGAGAGAGATGTGACAGAGAGAGAGAGAGAGATGTGACAGATTTGTTCC 2844
7723 AGGCAAGAGAGAGATGTGACAGAGAGAGAGAGAGAGATGTGACAGATTTGTTCC 7782
2845 TTGGTCTTGTGGAGCAGCAGAGAGAGATGTGAGGCTGACGATGACGCTGACGCTGAC 2904

7783 TTGGTCTTGTGGAGCAGCAGAGAGAGATGTGAGGCTGACGCTGACGCTGACGCTGAC 7842
2905 AGGCCAGCAATTTATTTCTGATATGTAGTACAGCAGCAGACAGATTTGCTGAGGCTATTG 2964
7843 AGGCCAGCAATTTATTTCTGATATGTAGTACAGCAGCAGACAGATTTGCTGAGGCTATTG 7902
2965 AGGCCAGCAGATTTGTTGACATCAGATCTGGGGATCAACAGCTCCAGAGAGAGAA 3024
7903 AGGCCAGCAGATTTGTTGACATCAGATCTGGGGATCAACAGCTCCAGAGAGAGAA 7962
3025 TCTGCTGCTGAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3084
7963 TCTGCTGCTGAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 8022
3085 GAAACATATTTGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3144
8023 GAAACATATTTGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8082
3145 AACAGATTTGAGATTAACATGACCTGATGAGATGGGACAGAGAGATTTAACAATTACAGA 3204
8083 AACAGATTTGAGATTAACATGACCTGATGAGATGGGACAGAGAGATTTAACAATTACAGA 8142
3205 GCTTATACACTCTCTAATTTGAAGATGCAAAAACAGAGAGAGAGATTAACAAGAT 3264
8143 GCTTATACACTCTCTAATTTGAAGATGCAAAAACAGAGAGAGATTTAACAAGAT 8202
3265 TATGGAATTTAGATTAATGAGCAGTGTGGAATTTGTTAACAATTAACAATTTGCTGT 3324
8203 TATGGAATTTAGATTAATGAGCAGTGTGGAATTTGTTAACAATTAACAATTTGCTGT 8262
3325 GGTATATTAATTTATTTCTAATGATAGTAGAGAGCTTGGTATGATTAAGATTTGTTT 3384
8263 GGTATATTAATTTATTTCTAATGATAGTAGAGAGCTTGGTATGATTAAGATTTGTTT 8322
3385 CTGTACTTCTATATGATAGATAGATTTAGAGAGATTTACCATTTATGTTTACAGACC 3444
8323 CTGTACTTCTATATGATAGATAGATTTAGAGAGATTTACCATTTATGTTTACAGACC 8382
3445 ACCTCCCAATCCCGAGGGAGCCGACAGAGCCGAGAGAGATTAAGAGAGAGATTTGAGAGA 3504
8383 ACCTCCCAATCCCGAGGGAGCCGACAGAGCCGAGAGAGATTAAGAGAGAGATTTGAGAGA 8442
3505 GAGACAGAGAGATTTGATGATAGTATGAGAGAGATTTGATGATGATGATGATGATGATG 3564
8443 GAGACAGAGAGATTTGATGATAGTATGAGAGAGATTTGATGATGATGATGATGATGATG 8502
3565 TGCGAGGCTGTGCTCTTCAAGCTACACCGCTTGAAGAGATTTGATGATGATGATGATGATG 3624
8503 TGCGAGGCTGTGCTCTTCAAGCTACACCGCTTGAAGAGATTTGATGATGATGATGATGATG 8562
3625 GGATTTGAGAACTTGTGAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3684
8563 GGATTTGAGAACTTGTGAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8622
3685 TACAGTATTTGAGATGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3744
8623 TACAGTATTTGAGATGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3804
3745 TAGCAGTATGAGAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3804
8623 TAGCAGTATGAGAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 8742
3805 TTGCGCATATCTTGAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3864
8743 TTGCGCATATCTTGAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 8802
3865 GGCAGAGAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3924
8803 GGCAGAGAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8862
3925 GCTGAGCAGCAGCAGATGAGGCTGAGCAGATTTCTGAGATTTCTGAGATTTCTGAGAT 3971

Db 6365 ATATCAGCACAAGATAGATAGAGTGCAGAAAAAATATGATCTTTTATTAACCTTG 6424
 QY 1831 ATATAGTCAATAGATATATACAGCTATAGTTGATTAAGTTGTACACCTCAGCATTA 1890
 Db 6425 ATATAGTCAATAGATATATACAGCTATAGTTGATTAAGTTGTACACCTCAGCATTA 6484
 QY 1891 CACAGGCGCTGCACAAAGGTATCTTTGAGCCCAATTCCTATCATTTATGTGCCGCGCTG 1950
 Db 6485 CACAGGCGCTGCACAAAGGTATCTTTGAGCCCAATTCCTATCATTTATGTGCCGCGCTG 6544
 QY 1951 GTTTTGGCATCTTCAAAATGTATATTAAGACGTTCAATGGAACGAGCATTGTACAAATG 2010
 Db 6545 GTTTTGGCATCTTCAAAATGTATATTAAGACGTTCAATGGAACGAGCATTGTACAAATG 6604
 QY 2011 TCACACAGTCAATGTACATGTAATGGAATCAGGCCAGTATGATCAACTCACTGCTGTAA 2070
 Db 6605 TCACACAGTCAATGTACATGTAATGGAATCAGGCCAGTATGATCAACTCACTGCTGTAA 6664
 QY 2071 ATGGCAGTCTAGCAGAAAGATGTAGTAATTTAGATCTGCCAATTTTCACAGCAATGCTA 2130
 Db 6665 ATGGCAGTCTAGCAGAAAGATGTAGTAATTTAGATCTGCCAATTTTCACAGCAATGCTA 6724
 QY 2131 AAACCATATATGTACAGCTGACATCTGTAGAATTAATTGTACAAAGACCAACA 2190
 Db 6725 AAACCATATATGTACAGCTGACATCTGTAGAATTAATTGTACAAAGACCAACA 6784
 QY 2191 ATACAGAAAGATATCCGATCCAGAGGGAGCAGGAGAGCATTTGTCAATAGGAA 2250
 Db 6785 ATACAGAAAGATATCCGATCCGATCCAGAGGGAGCAGGAGAGCATTTGTCAATAGGAA 6844
 QY 2251 AAATAGGAATATAGACACAGCATTGTAACTTGTAGAGCAAAATGGAATGCCATT 2310
 Db 6845 AAATAGGAATATAGACACAGCATTGTAACTTGTAGAGCAAAATGGAATGCCATT 6904
 QY 2311 TAAACAGATGTAGCAAAATTAAGAGACAATTTGGAATTAATAAATTAATTAATCTTA 2370
 Db 6905 TAAACAGATGTAGCAAAATTAAGAGACAATTTGGAATTAATAAATTAATTAATCTTA 6964
 QY 2371 AGCAATCCTCAGAGGGAGCCAGAAATTTGTACGACAGTTTAAATTTGGAGGGGAA 2430
 Db 6965 AGCAATCCTCAGAGGGAGCCAGAAATTTGTACGACAGTTTAAATTTGGAGGGGAA 7024
 QY 2431 TTTTCTACTGTAAATCAACACACTGTTTAACTACTGTTTAAATAGTCTTGGAGTA 2490
 Db 7025 TTTTCTACTGTAAATCAACACACTGTTTAACTACTGTTTAAATAGTCTTGGAGTA 7084
 QY 2491 CTGAAGGCTCAAAATTAACCTGAAGAGTGCACAACTCCATCCCATCAGTGGAC 2550
 Db 7085 CTGAAGGCTCAAAATTAACCTGAAGAGTGCACAACTCCATCCCATCAGTGGAC 7144
 QY 2551 AATTTATPAACATGTGCGAGAGATAGAAAAGCAATGTATGCCCTCCCATCAGTGGAC 2610
 Db 7145 AATTTATPAACATGTGCGAGAGATAGAAAAGCAATGTATGCCCTCCCATCAGTGGAC 7204
 QY 2611 AAATTAATGTTCATCAAAATTTACTGGGCTGCTATTAACAAAGATGTGTAAATACA 2670
 Db 7205 AAATTAATGTTCATCAAAATTTACTGGGCTGCTATTAACAAAGATGTGTAAATACA 7264
 QY 2671 ACAATGGGTCGAGATCTTTCAGACCTGAGAGGCGATATGAGGCAAAATTTGGAGAGTG 2730
 Db 7265 ACAATGGGTCGAGATCTTTCAGACCTGAGAGGCGATATGAGGCAAAATTTGGAGAGTG 7324
 QY 2731 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2790
 Db 7325 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7384
 QY 2791 AGAGAAAGTGTGTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2850
 Db 7385 AGAGAAAGTGTGTCAGAGAAAG 7444
 QY 2851 TCTTGGAG 2910
 Db 7445 TCTTGGAG 7504

QY 2911 GACATTAATGTCTGATATAGTGCAGAGCAAGCAATTTGCGAGGGCTATTGAGGCGC 2970
 Db 7505 GACATTAATGTCTGATATAGTGCAGAGCAAGCAATTTGCGAGGGCTATTGAGGCGC 7564
 QY 2971 AACGACATCTTGTCAACCTCAGACCTGTGGGGCATCAACAGCTCCAGGCAAGAAATCTG 3030
 Db 7565 AACGACATCTTGTCAACCTCAGACCTGTGGGGCATCAACAGCTCCAGGCAAGAAATCTG 7624
 QY 3031 CTGTGGAAGATTAATTAAGATTAACAGCTCTCTGGGGATTTGGGGTTGCTGTGAAAAC 3090
 Db 7625 CTGTGGAAGATTAATTAAGATTAACAGCTCTCTGGGGATTTGGGGTTGCTGTGAAAAC 7684
 QY 3091 TCATTTGCACCACTGCTGCTTGGAAATGCTAGTTGAGATTAATAATCTGTGAAACGA 3150
 Db 7685 TCATTTGCACCACTGCTGCTTGGAAATGCTAGTTGAGATTAATAATCTGTGAAACGA 7744
 QY 3151 TTTGGAATTAACATGACCTGTGAGTGGAGCAGAGAAATTAACAAATTAACAGCTTAA 3210
 Db 7745 TTTGGAATTAACATGACCTGTGAGTGGAGCAGAGAAATTAACAAATTAACAGCTTAA 7804
 QY 3211 TACACTCCTTAATTAAGAAATCCGAAAACAGCAAGAAAGAAATGAACAGAAATTAATGG 3270
 Db 7805 TACACTCCTTAATTAAGAAATCCGAAAACAGCAAGAAAGAAATGAACAGAAATTAATGG 7864
 QY 3271 AATTAGATTAATGAGGCAAGTTTGTGAATTTGTTTAACTAATTAACAAATTTGCTGTATA 3330
 Db 7865 AATTAGATTAATGAGGCAAGTTTGTGAATTTGTTTAACTAATTAACAAATTTGCTGTATA 7924
 QY 3331 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3390
 Db 7925 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7984
 QY 3391 TTTCTATAGGAATAGAGTTAGAGGAGATTAATCAGATTAATGCTTACAGCCACTCC 3450
 Db 7985 TTTCTATAGGAATAGAGTTAGAGGAGATTAATCAGATTAATGCTTACAGCCACTCC 8044
 QY 3451 CAATCCCGAGGGAGCCCGACAGGCGCCGAAGGAATGAAGAAAGAGTGTGAGAGAGAGACA 3510
 Db 8045 CAATCCCGAGGGAGCCCGACAGGCGCCGAAGGAATGAAGAAAGAGTGTGAGAGAGAGACA 8104
 QY 3511 GAGACAGATCCATTCGATTAGTGAAGCGATCTTGAAGCACTTATCTGGAGAGATCTCGGGA 3570
 Db 8105 GAGACAGATCCATTCGATTAGTGAAGCGATCTTGAAGCACTTATCTGGAGAGATCTCGGGA 8164
 QY 3571 GCCGTGCTCTTACAGTACACCGCTTGGAGAGCTTACTGATTTGAACGAGAGTTG 3630
 Db 8165 GCCGTGCTCTTACAGTACACCGCTTGGAGAGCTTACTGATTTGAACGAGAGTTG 8224
 QY 3631 TGGAACTTCTGGGAGCGAGGGGTGGGAAGCCCTCAAAATTAATGTTGGAATCTCTACAGT 3690
 Db 8225 TGGAACTTCTGGGAGCGAGGGGTGGGAAGCCCTCAAAATTAATGTTGGAATCTCTACAGT 8284
 QY 3691 ATTGGAATCAGGAATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3750
 Db 8285 ATTGGAATCAGGAATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8344
 QY 3751 TAGCTGAGGAGCAGATAGGGTTATAGAAATTAACAGAGCTTATAGAGCTATTCCGC 3810
 Db 8345 TAGCTGAGGAGCAGATAGGGTTATAGAAATTAACAGAGCTTATAGAGCTATTCCGC 8404
 QY 3811 ACATACCTAGAAAGATTAAGACAGGGCTTGAAGAGATTTTGTATTAAGATGGGTGCAAG 3870
 Db 8405 ACATACCTAGAAAGATTAAGACAGGGCTTGAAGAGATTTTGTATTAAGATGGGTGCAAG 3930
 QY 3871 TGGT 3874
 Db 8465 TGGT 8468

RESULT 15
 US-08-935-312-13
 ; Sequence 13, Application US/08935312

Patent No. 6207455
GENERAL INFORMATION:
APPLICANT: CHANG, Lung-Ji
TITLE OF INVENTION: LENTIVIRAL VECTORS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P. L. L. C.
STREET: 624 Ninth Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/935,312
FILING DATE: 22-SEP-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: CHANG-112
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 12494 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
US-08-935-312-13

Query Match 41.1%; Score 2558.4; DB 3; Length 12494;
Best Local Similarity 99.4%; Pred. No. 0; Mismatches 16; Indels 0; Gaps 0;
Matches 2568; Conservative 0;

QY 1291 CCATGAGAGTGAAGAGAGATATCAGCATTGTGAGATGGGGTGGAAATGGGGCAACA 1350
DB 5890 CAATGAGAGTGAAGAGAGATATCAGCATTGTGAGATGGGGTGGAAATGGGGCAACA 5949
QY 1351 TGTCTCTTGGGATATGTGATGATCTGTAGTGTACAGAAAAATGTGGGTCAACGCTTATT 1410
DB 5950 TGTCTCTTGGGATATGTGATGATCTGTAGTGTACAGAAAAATGTGGGTCAACGCTTATT 6009
QY 1411 ATGGGGTACCTGTGTGGAAGAACCAACCACTCTATTGTCATCAATGCTTAAG 1470
DB 6010 ATGGGGTACCTGTGTGGAAGAACCAACCACTCTATTGTCATCAATGCTTAAG 6069
QY 1471 CATATGATACAGAGTACATATGTTGGGCGACACATGCTGTACCCACAGACCCA 1530
DB 6070 CATATGATACAGAGTACATATGTTGGGCGACACATGCTGTACCCACAGACCCA 6129
QY 1531 ACCCACAAGAGTAGTATGTTGTAATGTGACAGAAAAATTTTAAACATGTGAAAAATGACA 1590
DB 6130 ACCCACAAGAGTAGTATGTTGTAATGTGACAGAAAAATTTTAAACATGTGAAAAATGACA 6189
QY 1591 TGTGTAGAAGCATGATGAGATATTAATCACTTATGAGATCAAAAGCCTAAAGCATGCTG 1650
DB 6190 TGTGTAGAAGCATGATGAGATATTAATCACTTATGAGATCAAAAGCCTAAAGCATGCTG 6249
QY 1651 TAAATTAACCCCACTGTGTAGTTTAAAGTGCACATGATTTGAAGATGATCTATA 1710
DB 6250 TAAATTAACCCCACTGTGTAGTTTAAAGTGCACATGATTTGAAGATGATCTATA 6309
QY 1711 CCAATAGTAGTAGCGGGAGATGATTAATGAGAAAGAGAGATAAAAATGCTCTTTCA 1770
DB 6310 CCAATAGTAGTAGCGGGAGATGATTAATGAGAAAGAGAGATAAAAATGCTCTTTCA 6369

QY 1771 ATATCAGACAAAGCATAGATTAAGTGCAGAAAGAAATATGATCTTTTATAAAGCTTG 1830
DB 6370 ATATCAGACAAAGCATAGATTAAGTGCAGAAAGAAATATGATCTTTTATAAAGCTTG 6429
QY 1831 ATATAGTACCAATAGATTAATACAGCTATAGTTGATTAAGTTGTAAACCTCACTCATTA 1890
DB 6430 ATATAGTACCAATAGATTAATACAGCTATAGTTGATTAAGTTGTAAACCTCACTCATTA 6489
QY 1891 CACAGGCTGTCCAAAGGATATCTTTGAGCCAAATCCCATTCATTTATTTGCCCCGGCTG 1950
DB 6490 CACAGGCTGTCCAAAGGATATCTTTGAGCCAAATCCCATTCATTTATTTGCCCCGGCTG 6549
QY 1951 GTTTTGGATCTTAAATGTAAATTAAGAGCTTCAATGGAACGAGACATGTACAAATG 2010
DB 6550 GTTTTGGATCTTAAATGTAAATTAAGAGCTTCAATGGAACGAGACATGTACAAATG 6609
QY 2011 TCAGCAGATCAATGTACACATGGAATCAGGCCAGTAGTATCAACTCACTGCTTTAA 2070
DB 6610 TCAGCAGATCAATGTACACATGGAATCAGGCCAGTAGTATCAACTCACTGCTTTAA 6669
QY 2071 ATGCGAGCTTAGCAGAGAGATGATGATTAATGATCTGCCAATTTTACAGACATGCTA 2130
DB 6670 ATGCGAGCTTAGCAGAGAGATGATGATTAATGATCTGCCAATTTTACAGACATGCTA 6729
QY 2131 AAACCATATATGTACAGCTGAACACATCTGTAGAATTAATGTACAAAGACCAACACA 2190
DB 6730 AAACCATATATGTACAGCTGAACACATCTGTAGAATTAATGTACAAAGACCAACACA 6789
QY 2191 ATACAGAAAAAGTATCCGTATCCAGAGGGGAGCAGAGAGCATTTGTTCAATAGGAA 2250
DB 6790 ATACAGAAAAAGTATCCGTATCCAGAGGGGAGCAGAGAGCATTTGTTCAATAGGAA 6849
QY 2251 AAATAGGAATATATGAGACACACATTTTAACATTAGTAGAGCAAAATGATGCACTT 2310
DB 6850 AAATAGGAATATATGAGACACACATTTTAACATTAGTAGAGCAAAATGATGCACTT 6909
QY 2311 TAAACAGATAGCTAGCAAAATTAAGAGAAATTTGGAATTAATTAACATTAATCTTTA 2370
DB 6910 TAAACAGATAGCTAGCAAAATTAAGAGAAATTTGGAATTAATTAACATTAATCTTTA 6969
QY 2371 AGCAATCTCTCAGAGGGGAGCCAGAAATTTGTAAAGCAGCTTTTAATTTGTGAGGGGAT 2430
DB 6970 AGCAATCTCTCAGAGGGGAGCCAGAAATTTGTAAAGCAGCTTTTAATTTGTGAGGGGAT 7029
QY 2431 TTTTCTACTGTAAATTCACACACACTTTTAATAGTACTTGTGTTTAATAGTCTTGAGCTA 2490
DB 7030 TTTTCTACTGTAAATTCACACACACTTTTAATAGTACTTGTGTTTAATAGTCTTGAGCTA 7089
QY 2491 CTGAAGGCTCAAAATTAACCTGAAGAGTGAACACATCCACTCCATGCAAGTAATTAAC 2550
DB 7090 CTGAAGGCTCAAAATTAACCTGAAGAGTGAACACATCCACTCCATGCAAGTAATTAAC 7149
QY 2551 AATTTATTAACATGTGCGAGAGAGTAGGAAAGCAATGATCCCTCCATCACTAGTGAGC 2610
DB 7150 AATTTATTAACATGTGCGAGAGAGTAGGAAAGCAATGATCCCTCCATCACTAGTGAGC 7209
QY 2611 AAATTTAGTGTTCATCAATATTTACTGGGCTGCTATTTAACAGAGATGCTGTAATACA 2670
DB 7210 AAATTTAGTGTTCATCAATATTTACTGGGCTGCTATTTAACAGAGATGCTGTAATACA 7269
QY 2671 ACAATGGTCCGAGATCTTACAGACTGGAAGAGGAGATATAGGAGCAATTTGGAGAAATG 2730
DB 7270 ACAATGGTCCGAGATCTTACAGACTGGAAGAGGAGATATAGGAGCAATTTGGAGAAATG 7329
QY 2731 AATTATATTAATTAATAGTAGTAAATTTGAACATTTAGAGTAGAGCACCACCAAGGCAA 2790
DB 7330 AATTATATTAATTAATAGTAGTAAATTTGAACATTTAGAGTAGAGCACCACCAAGGCAA 7389
QY 2791 AGAGAAAGTGTGTCAGAGAGAGAAAAAGAGCAGTGGGAATAGAGCTTTGTTCTTGCGT 2850
DB 7390 AGAGAAAGTGTGTCAGAGAGAGAAAAAGAGCAGTGGGAATAGAGCTTTGTTCTTGCGT 7449

Search completed: September 17, 2003, 17:30:45
Job time : 337 secs

```
QY 2851 TCTTGGAGCAGCAGGAAGCACTATGCGCTGCAGTCAATGACGCTGACGGGTACAGGCCA 2910
Db 7450 TCTTGGAGGAGCAGGAAGCACTATGCGCGACGCTCATATACGCTGACGGGTACAGGCCA. 7509
QY 2911 GACAATTATTTGTCGATATAGTCAGCAGCAGAACAAATTTGCTGAGGCTATTTGAGCGCC 2970
Db 7510 GACAATTATTTGTCGATATAGTCAGCAGCAGAACAAATTTGCTGAGGCTATTTGAGCGCC 7569
QY 2971 AACAGCATCTGTGCAACTCAGACGCTGCGGCAATCAACACGCTCCAGGCAAGAAATCTCG 3030
Db 7570 AACAGCATCTGTGCAACTCAGACGCTGCGGCAATCAACACGCTCCAGGCAAGAAATCTCG 7629
QY 3031 CTGTGGAAGAGATACCTAAAGGATCAACAGCTCTGTGGGATTTGGGGTTCCTGTGAAAAC 3090
Db 7630 CTGTGGAAGAGATACCTAAAGGATCAACAGCTCTGTGGGATTTGGGGTTCCTGTGAAAAC 7689
QY 3091 TCATTTGCACACGCTGCTGCTGTGGAATGCTATGTGAGTAATTAATCTGTGAGACAGA 3150
Db 7690 TCATTTGCACACGCTGCTGCTGTGGAATGCTATGTGAGTAATTAATCTGTGAGACAGA 7749
QY 3151 TTTGGAATTAACATGACCTGATGAGTGAGAGAGAGAAATTAACAAATTAACAAAGCTTAA 3210
Db 7750 TTTGGAATTAACATGACCTGATGAGTGAGAGAGAGAAATTAACAAATTAACAAAGCTTAA 7809
QY 3211 TACACTCCTTAATTGAAGATTCGCAAAACAGCAAGAAAAGATGAACAGAAATTTATGG 3270
Db 7810 TACACTCCTTAATTGAAGATTCGCAAAACAGCAAGAAAAGATGAACAGAAATTTATGG 7869
QY 3271 AATTAGATTAATGGGCAAGTTTGGAAATTTGGTTAACTAACAAATTTGGCTGTGTATA 3330
Db 7870 AATTAGATTAATGGGCAAGTTTGGAAATTTGGTTAACTAACAAATTTGGCTGTGTATA 7929
QY 3331 TAAATTTATTCATATGATAGTAGAGGCTTGTAGTTTAAAGATAGTTTGTCTGTAC 3390
Db 7930 TAAATTTATTCATATGATAGTAGAGGCTTGTAGTTTAAAGATAGTTTGTCTGTAC 7989
QY 3391 TTTCTATAGTGAATAGATTAGCAGGAGATTTCAACATTAATTCGTTTCAGACCACTCC 3450
Db 7990 TTTCTATAGTGAATAGATTAGCAGGAGATTTCAACATTAATTCGTTTCAGACCACTCC 8049
QY 3451 CAATCCGAGGGGAGCCGACAGGCGCCGAGGAATAGAAAGAGTGGAGAGAGACA 3510
Db 8050 CAATCCGAGGGGAGCCGACAGGCGCCGAGGAATAGAAAGAGTGGAGAGAGACA 8109
QY 3511 GAGACAGATCCATTGATTAAGTGAACGATCCCTTAGCACTTATCTGGGACGATCTGCGGA 3570
Db 8110 GAGACAGATCCATTGATTAAGTGAACGATCCCTTAGCACTTATCTGGGACGATCTGCGGA 8169
QY 3571 GCCGTGCTCTTACGCTACACCGCTTGAGAGACTTACTGTTGATTTGAAGAGATTG 3630
Db 8170 GCCGTGCTCTTACGCTACACCGCTTGAGAGACTTACTGTTGATTTGAAGAGATTG 8229
QY 3631 TGSAACTTTGGGAGCGAGGGGGTGGGAAGCCCAATATTTGGGGAATCTCTACAGT 3690
Db 8230 TGSAACTTTGGGAGCGAGGGGGTGGGAAGCCCAATATTTGGGGAATCTCTACAGT 8289
QY 3691 ATTGAGTCAAGAACTAAAGATAGTCTGTTAACTTGCATGCAATGCACAGCCATAGCAG 3750
Db 8290 ATTGAGTCAAGAACTAAAGATAGTCTGTTAACTTGCATGCAATGCACAGCCATAGCAG 8349
QY 3751 TAGCTGAGGGAGCAGATAGGGTTATAGAGTATTACAGACGCTTATAGAGCTATTGCCC 3810
Db 8350 TAGCTGAGGGAGCAGATAGGGTTATAGAGTATTACAGACGCTTATAGAGCTATTGCCC 8409
QY 3811 ACATACCTAGAAATTAAGACAGGGCTTGAAGAGATTTGCTTAAGATGGGTGCAAG 3870
Db 8410 ACATACCTAGAAATTAAGACAGGGCTTGAAGAGATTTGCTTAAGATGGGTGCAAG 8469
QY 3871 TGGT 3874
Db 8470 TGGT 8473
```


Db	6582	CATATGATACAGAGGTACATATATGTTGGCCACACATGCTGTGTACCCACAGACCCCA	66411
QY	1531	ACCCACAGAAGTAGTATTTGGTAAATGTGACAGAAAAATTTTAACTGTGTGAAAAATGACA	15900
Db	6642	ACCCACAGAAGTAGTATTTGGTAAATGTGACAGAAAAATTTTAACTGTGTGAAAAATGACA	67010
QY	1591	TGTGTAGACACATGCATGAGTATGATTAACATTTTATGGGATCCAAAGCCTTAAGCCATGTG	16503
Db	6702	TGTGTAGAACACATGCATGAGTATTAACATTTTATGGGATCCAAAGCCTTAAGCCATGTG	67610
QY	1651	TAAATTTAAACCCACTGTGTGTATGTTTAAAGTGACACGATTTTGAAGTAATGATACTAATA	17110
Db	6762	TAAATTTAAACCCACTGTGTGTATGTTTAAAGTGACACGATTTTGAAGTAATGATACTAATA	68210
QY	1711	CCAAATAGTAGTAGCGGGGAGAAATGATATGAGAAAAAGAGATATAAAAACTGCTCTTTTCA	17700
Db	6822	CCAAATAGTAGTAGCGGGGAGAAATGATATGAGAAAAAGAGATATAAAAACTGCTCTTTTCA	68810
QY	1771	ATATTCAGACACAGCATTAAGAGATTAAGGTGCGAGAAAAGTAATATCATTTCTTTTATAAATCTG	18300
Db	6882	ATATTCAGACACAGCATTAAGAGGTAAAGGTGCGAGAAAAGTAATATCATTTCTTTTATAAATCTG	69410
QY	1831	ATATAGTATCCATAGATATA-----TACACAGTATAGTGTGATTAAGTTGTATACACCTTAG	18840
Db	6942	ATATATATATCCATATAGATATATATATATCTACCAAGCTATATAGTTGTACAGATGTGTACACCTTAG	70010
QY	1885	TCATTTTACACAGGCGCTGTCCAAAGGTATTCCTTTTGAGCCAAATCCCATCATTTATTTGTGCC	19440
Db	7002	TCATTTTACACAGGCGCTGTCCAAAGGTATTCCTTTTGAGCCAAATCCCATCATTTATTTGTGCC	70610
QY	1945	CGCGTGGTGTTCGCATTTTAAATATGTATTAATTAAGAGCTTCAATGGAACAGAGACCATGTA	20040
Db	7062	CGCGTGGTGTTCGCATTTTAAATATGTATTAATTAAGAGCTTCAATGGAACAGAGACCATGTA	71210
QY	2005	CAAAATGTGACACAGTACATGTACACATGTCATGATGACGCCGCGTAGTATCAACTCAATCCGC	20640
Db	7122	CAAAATGTGACACAGTACAAATGTACACATGTGCAATTTAGGCCGCGTAGTATCAACTCAACTGC	71810
QY	2065	TGTTTAAATGGAGTCTAGCAGAGAGAAGATGTATTTAGATTCGCCAAATTTTCACAGACA	21240
Db	7182	TGTTTAAATGGAGTCTAGCAGAGAGAAGATGTATTTAGATTCGCCAAATTTTCACAGACA	72410
QY	2125	ATGCTTAAACCATATATAGTACAGCTGAACACATCTGTGTGAATTTATTTGTACAGACCCA	21840
Db	7242	ATGCTTAAACCATATATAGTACAGCTGAACACATCTGTGTGAATTTATTTGTACAGACCCA	73010
QY	2185	ACAACAAATACAGAAAAAGTATCCGTATCCAGAGGGGACCCAGGGGAGACATTTGTTTACAA	22440
Db	7302	ACAACAAATACAGAAAAAGTATCCGTATCCAGAGGGGACCCAGGGGAGACATTTGTTTACAA	73610
QY	2245	TAGGAAAAATATGAGAAATATGAGACACAGCATTTGTAAATATTAGTGAAGCAAAATGGATG	23040
Db	7362	TAGGAAAAATATGAGAAATATGAGACACAGCATTTGTAAATATTAGTGAAGCAAAATGGATG	74210
QY	2305	CCACTTTTAAACAGATAGCTAGCAAAATTAAGAGAACAAATTTGGAATATATTAACAAATTA	23640
Db	7422	ACACTTTTAAACAGATAGCTAGCAAAATTAAGAGAACAAATTTGGAATATATTAACAAATTA	74810
QY	2365	TCTTTTAAACCATCTCAGAGAGGGACCCAGAAATTTGTAGCAGACAGTTTAAATTTGTGGAG	24240
Db	7482	TCTTTTAAACCATCTCAGAGAGGGACCCAGAAATTTGTAGCAGACAGTTTAAATTTGTGGAG	75410
QY	2425	GGGAATTTTCTTACGTATATCAACACACACGTTTAAATAGTACCTGTTTAAATAGTACTT	24840
Db	7542	GGGAATTTTCTTACGTATATCAACACACACGTTTAAATAGTACCTGTTTAAATAGTACTT	76010
QY	2485	GGAGTACTGAAGGGTCAATTAACACTGAAGAAAGTACACATATCCATATGTCAGAA	25440
Db	7602	GGAGTACTGAAGGGTCAATTAACACTGAAGAAAGTACACATATCCATATGTCAGAA	76610
QY	2545	TAAACAAATTTATAAACATGTGTGCGAGAGATAGGAAAAAGCAATGTATGCCCTCCATCA	26040

D	b	7662	TAAACAATATTTAAACATCTGGCGAAGAACTAGAGAAAAGCAATGTATGCCCTCCATCA	7721
Q	y	2605	GTGGCAAAATATAGATGTTTCATCAATATATTTACTGGCTGCTATTTAAACAAGAGATGGTGTA	2664
D	b	7722	GTGGCAAAATATAGATGTTTCATCAATATATTTACAGGGCTCTATTTAAACAAGATGGTGTA	7781
Q	y	2665	ATACCAACAATAGGGCCGAGATCTTCAGACTGGAGGGCGGATATGTGGGACAAATTGGA	2724
D	b	7782	ATAGCAACAATAGAGTCGAGATCTTCACACCTGGAGGGGAGATATGTGGGACAAATTGGA	7841
Q	y	2725	GAAGTAATTTATATAATATTAAGTAGTAATAAAATTTGAACCTATTAGAGATAGACCCACA	2784
D	b	7842	GAAGTAATTTATATAATATATAAGTAGTAATAAAATTTGAACCTATTAGAGATAGACCCACA	7901
Q	y	2785	AGGCAAGAGAAGAGCTGTCGACAGAGAAAAAGACAGTGGGAATVAGAGCTTTGTTCC	2844
D	b	7902	AGGCAAAAGAGAAGTGGTCGACAGAGAAAAAGACAGTGGGAATVAGAGCTTTGTTCC	7961
Q	y	2845	TTGGGTTCTTGGGAGACACAGGAAGCACTATGGGCTGCACCTCAATGACCTGACGGTAC	2904
D	b	7962	TTGGGTTCTTGGGAGACACAGGAAGCACTATGGGCGCAGCTTCATAGCAGCTGACGGTAC	8021
Q	y	2905	AGGCGACAATATTTATGCTATTTATGTCAGACAGACAGAAATTTGCTGAGGCTATTC	2964
D	b	8022	AGGCGACAATATTTATGCTGCTGTAATAGTCAGACAGACAGAAATTTGCTGAGGCTATTC	8081
Q	y	2965	AGGCGACAAGACATCTGTTGGCACTCACAGTCTGGGGCATCAACAGCTCCAGGCAAGAA	3024
D	b	8082	AGGCGACAAGACATCTGTTGGCACTCACAGTCTGGGGCATCAACAGCTCCAGGCAAGAA	8141
Q	y	3025	TCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCTCTGGGGATTTGGGGTTCCTTG	3084
D	b	8142	TCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCTCTGGGGATTTGGGGTTCCTTG	8201
Q	y	3085	GAAGAATATTTGACACCACTGCTGCTGGCTTGGAAATGCTAGTTGGAGTAATTAATCTCGG	3144
D	b	8202	GAAGAATATTTGACACCACTGCTGCTGGCTTGGAAATGCTAGTTGGAGTAATTAATCTCGG	8261
Q	y	3145	AACAGATTTGGAATPACATGACCTTGATGAGTGGGACAGAAATTTAACATTTACACAA	3204
D	b	8262	AACAGATTTGGAATPACATGACCTTGATGAGTGGGACAGAAATTTAACATTTACACAA	8321
Q	y	3205	GCTTAATACACTCCTTAATTTGAACAAATCGCAAAACCCAGAAAGAAATGAACAAGAT	3264
D	b	8322	GCTTAATACACTCCTTAATTTGAAGATTCGCAAAACCCAGCAAGAAAGATGAACAAGAT	8381
Q	y	3265	TATTGGAAATTTAGATTAATGGGCAAGTTGTGGAATTTGGTTTAACTAACAATTTGGCTGT	3324
D	b	8382	TATTGGAAATTTAGATTAATGGGCAAGTTGTGGAATTTGGTTTAACTAACAATTTGGCTGT	8441
Q	y	3325	GGTATATATAATTTATTCATTAATGATAGTAGAGGCTTGGTAGTTTGAAGATAGTTTGG	3384
D	b	8442	GGTATATATAATTTATTCATTAATGATAGTAGAGGCTTGGTAGTTTGAAGATAGTTTGG	8501
Q	y	3385	CTGTACTTTCTATAGTGAATAGAGTTAGGCAAGGATATTCACCAATTCGTTTCAGACCC	3444
D	b	8502	CTGTACTTTCTATAGTGAATAGAGTTAGGCAAGGATATTCACCAATTCGTTTCAGACCC	8561
Q	y	3445	ACCTTCCCAATCCCGAGGGGACCCGACAGGCCCCGAAGAGATAGAAAGAGAAGAGTGGAGGA	3504
D	b	8562	ACCTTCCCAACCCCGAGGGGACCCGACAGGCCCCGAAGAGATAGAAAGAGAAGAGTGGAGGA	8621
Q	y	3505	GAGACAGAGACAGATCCATTCGATTAGTGAACGGATCTTGGCACTTATCTGGGACGATC	3564
D	b	8622	GAGACAGAGACAGATCCATTCGATTAGTGAACGGATCTTGGCACTTATCTGGGACGATC	8681
Q	y	3565	TGCGAGAGCTGTGCTTTCAGCTAACCAACGCTTTGAGAGACTTCTTGTATTTGAACGA	3624
D	b	8682	TGCGAGAGCTGTGCTTTCAGCTAACCAACGCTTTGAGAGACTTCTTGTATTTGAACGA	8741
Q	y	3625	GGATTGTGGAACCTTCTGGAGCGCAGGGGGTGGGAACCTCTAATATTTGGTGAATCTCC	3684
D	b	8742	GGATTGTGGAACCTTCTGGAGCGCAGGGGGTGGGAACCTCTAATATTTGGTGAATCTCC	8801

OY	3685	TACAGTATTGGATCAGGACCTAAAGATAGTGTGTTAACTGCTCAATGCGCAGGCA	3744
Db	8802	TACAGTATTGGATCAGGACCTAAAGATAGTGTGTTAACTGCTCAATGCGCAGGCA	8861
OY	3745	TAGCAGTAGCTGAGGGGACAGATAGGCTTATAGAGTATTACAAAGCAGCTTATAGGCTA	3804
Db	8862	TAGCAGTAGCTGAGGGGACAGATAGGCTTATAGAGTATTACAAAGCAGCTTATAGGCTA	8921
OY	3805	TTCCGCACATACCTTGAAGATTAAGACAGGGCTTGGAAAGAGTTTGGCTATTAAGTGGCT	3864
Db	8922	TTCCGCACATACCTTGAAGATTAAGACAGGGCTTGGAAAGAGTTTGGCTATTAAGTGGCT	8981
OY	3865	GGCAAGTGGTAAAAAGTAGTGTGATGCTGCTGCTTAAGGGAAGATGAGACGA	3924
Db	8982	GGCAAGTGGTAAAAAGTAGTGTGATGCTGCTGCTTAAGGGAAGATGAGACGA	9041
OY	3925	GCTGAGCCAGACAGATGGGGTGGAGCAGTATCTGA	3963
Db	9042	GCTGAGCCAGACAGATGGGGTGGAGCAGTATCTGA	9101
OY	3964	TTATTTGACGTTTAAATGTTTACAAATAAGCAATAGATCACAATTTACAAATTAAG	9161
Db	9102	TTATTTGACGTTTAAATGTTTACAAATAAGCAATAGATCACAATTTACAAATTAAG	9161
OY	3964	CATTTTTTTCACTGATTCATGTTGTGTGTCCAAACCTCATCATGATCTTATCTATG	9221
Db	9162	CATTTTTTTCACTGATTCATGTTGTGTGTCCAAACCTCATCATGATCTTATCTATG	9221
OY	3964	TCGTGTCGACTGTAGAGCTAAGCAAGTATGCAAAATGCAAGTGTGCTTCGATTTGGAAA	9281
Db	9282	TCGTGTCGACTGTAGAGCTAAGCAAGTATGCAAAATGCAAGTGTGCTTCGATTTGGAAA	9281
OY	3964	TTCCCAATCTGGTGGCCAAAGAGGTTCAATAATACAGTAATGGAACCTCTTATGTCAAAAGT	9341
Db	9282	TTCCCAATCTGGTGGCCAAAGAGGTTCAATAATACAGTAATGGAACCTCTTATGTCAAAAGT	9341
OY	3964	GAGTACTGTTTAAAGAGCTAGATTGGAAACAGCTTCAGAGATGCTTTGGACGTGCGATGTGT	9401
Db	9342	GAGTACTGTTTAAAGAGCTAGATTGGAAACAGCTTCAGAGATGCTTTGGACGTGCGATGTGT	9401
OY	3963	TTTCATCCAGACATGATTAAGTATACATTGATGATTTGGACAAACCA	4028
Db	9402	TTTCATCCAGACATGATTAAGTATACATTGATGATTTGGACAAACCA	4028
OY	4029	TTCCACTGCAAGACATTCACAACTTCATGATGATGTTGTCACACAGGCTCCTGGAAGGC	9461
Db	9402	TTCCACTGCAAGACATTCACAACTTCATGATGATGTTGTCACACAGGCTCCTGGAAGGC	9461
OY	4029	AACTTACTTCACTGCTGCAAAAGTGAACCTAATAAAGGAATGGTTGTCATCATGCGCCGTGAT	9521
Db	9462	AACTTACTTCACTGCTGCAAAAGTGAACCTAATAAAGGAATGGTTGTCATCATGCGCCGTGAT	9521
OY	4029	GCCACTGACATTAAGAAAGCAATGACAAGAGCGATGAAGTATTTGGTGGCAGATGG	9581
Db	9522	GCCACTGACATTAAGAAAGCAATGACAAGAGCGATGAAGTATTTGGTGGCAGATGG	9581
OY	4029	CACTAGATATGCACTGAGTGAAGAAATGCTTTATTTTGAAT	4068
Db	9582	CACTAGATATGCACTGAGTGAAGAAATGCTTTATTTTGAAT	4068
OY	4069	TTTGATGCTAATGCTTATTTGTTGATTAATAGCTGCAATTAAGCAAGTAAACAA	4128
Db	9642	TTTGATGCTAATGCTTATTTGTTGATTAATAGCTGCAATTAAGCAAGTAAACAA	4128
OY	4129	CAATTCATTCATTTTATGTTTCAAGTT	4156
Db	9702	CAATTCATTCATTTTATATATTTGTTGTTGTTAATTCATGACAGTATTTATTTATAGT	9761
OY	4157	CAATTCATTCATTTTATGTTTCAAGTT	4159
Db	9762	CAATTCATTCATTTTATGTTTCAAGTT	9821
OY	4160	GGGGAGCTGGGAGGTTTAAAGCAAGTAAACCTCTCAAAATGTTGATGGCTGAT	4219
Db	9822	GGGGAGCTGGGAGGTTTAAAGCAAGTAAACCTCTCAAAATGTTGATGGCTGAT	9881

OY	4220	TATGATCT	4228
Db	9882	TATGATCT	9941
OY	4229	TCGTGAACTAAGTAGTATAATACCTTCTGACTTGTGAGAAATGTGAATACTAGATC	10001
Db	9942	TCGTGAACTAAGTAGTATAATACCTTCTGACTTGTGAGAAATGTGAATACTAGATC	10001
OY	4229	GCCTGCGCGCTTTGGGTGATACAGGTGAAGAAACCTGACATGACAGCTCCG	4281
Db	10002	GCCTGCGCGCTTTGGGTGATACAGGTGAAGAAACCTGACATGACAGCTCCG	10061
OY	4282	G-----AGACGCTGACAGCTTGTGTGAAGCGGATGCGCGAGCAG	4322
Db	10062	G-----AGACGCTGACAGCTTGTGTGAAGCGGATGCGCGAGCAG	10121
OY	4333	A-----CAAGCCGCTC	4333
Db	10122	A-----CAAGCCGCTC	10181
OY	4334	AGGGCGCTGACAGCGGCTTTGGCGGGTGTGCGGGGCGAGCATGACCGACAGTACG	4393
Db	10182	AGGGCGCTGACAGCGGCTTTGGCGGGTGTGCGGGGCGAGCATGACCGACAGTACG	10241
OY	4394	ATAGCGAGTGTATCTGCTTAA	4417
Db	10242	ATAGCGAGTGTATCTGCTTAA	10301
OY	4418	CGCGGGAGAGCGGCTTGGCATTTGGGCGCTTCCGCTTCCTGCTACGATCGCT	4417
Db	10302	CGCGGGAGAGCGGCTTGGCATTTGGGCGCTTCCGCTTCCTGCTACGATCGCT	10361
OY	4418	CTATGCGGCTATCAGACAGATTTACTGAGAGTGCACCATATGTC	4462
Db	10362	CTATGCGGCTATCAGACAGATTTACTGAGAGTGCACCATATGTC	10421
OY	4463	GGGCGGCTTGGTGGCGTTTTCATAGGCTCCGCCCCCTGACGA	4508
Db	10422	GGGCGGCTTGGTGGCGTTTTCATAGGCTCCGCCCCCTGACGA	10481
OY	4509	GCATCACAATAATGACGCTCAAGTCAGAGTGGCGAAACCCGACAGCATTAAGATA	4568
Db	10482	GCATCACAATAATGACGCTCAAGTCAGAGTGGCGAAACCCGACAGCATTAAGATA	10541
OY	4569	CGAGGCTTTCCCTGGAGAGTCCCTGCGGCTCTCCTGTCGACCTGCGCCCTTAC	4628
Db	10542	CGAGGCTTTCCCTGGAGAGTCCCTGCGGCTCTCCTGTCGACCTGCGCCCTTAC	10601
OY	4629	CGGATACCTGTCCGCTTCTCCCTTCCGGAAGCTGGCGCTTCTCATAGCTCACGCTG	4688
Db	10602	CGGATACCTGTCCGCTTCTCCCTTCCGGAAGCTGGCGCTTCTCATAGCTCACGCTG	10661
OY	4689	TAGGATCTCAGTTGCGGTGATGTCGCTCCAGCTGGGCTGTGTCAGAAACCCC	4748
Db	10662	TAGGATCTCAGTTGCGGTGATGTCGCTCCAGCTGGGCTGTGTCAGAAACCCC	10721
OY	4749	CGTTAGGCGCGACGCTGCGCTTATCCGGTAACTATGCTTGAAGTCAACCCGCTAAG	4808
Db	10722	CGTTAGGCGCGACGCTGCGCTTATCCGGTAACTATGCTTGAAGTCAACCCGCTAAG	10781
OY	4809	ACACGACTTATGCGCACTGGCAGCAGCACTGTGAAGATTTAGCAGAGCGAGTATGT	4868
Db	10782	ACACGACTTATGCGCACTGGCAGCAGCACTGTGAAGATTTAGCAGAGCGAGTATGT	10841
OY	4869	AGGCGTCTCAGAGTCTTGAAGTGTGGCTTAACCTAAGGCTACATAGAGACAGT	4928
Db	10842	AGGCGTCTCAGAGTCTTGAAGTGTGGCTTAACCTAAGGCTACATAGAGACAGT	10901
OY	4929	ATTGATGATTCGCGCTGCTGGAAGCCAGTTACCTTGGAAAAAGAGTTGGATGCTTG	4988
Db	10902	ATTGATGATTCGCGCTGCTGGAAGCCAGTTACCTTGGAAAAAGAGTTGGATGCTTG	10961
OY	4989	ATCCGCAAAACCAACCGCTGTAGCGGTGTTTTTTTGTGTCAGACAGATTTAC	5048

```
|||||
Db 10962 ATCCGGAACAAACCCCGCTGAGCGGCTTTTGTGCAAGCAGCATATAC 11021
Oy 5049 GCGGAGAAAAAGATCTCAAGAAATCCTTGTATCTTCTACGGGGGTGACGGCA 5108
Db 11022 GCGGAGAAAAAGATCTCAAGAAATCCTTGTATCTTCTACGGGGGTGACGGCA 11081
Oy 5109 GTGAGACGAAACACACAGTTAAGGATTTGGTCATGATGATTAACAAAAGATCTTAC 5168
Db 11082 GTGAGACGAAACACACAGTTAAGGATTTGGTCATGATGATTAACAAAAGATCTTAC 11141
Oy 5169 CTGATCCTTTTAAATTAATTAATGAAGTTTAAATCAATCAATTAATATATGATTAAC 5228
Db 11142 CTGATCCTTTTAAATTAATTAATGAAGTTTAAATCAATCAATTAATATATGATTAAC 11201
Oy 5229 TTGGTCGACGATTAACAAAGCTTAATCAGTACGAGGACCTTCTCAGGATCTCTAT 5288
Db 11202 TTGGTCGACGATTAACAAAGCTTAATCAGTACGAGGACCTTCTCAGGATCTCTAT 11261
Oy 5289 TCGTTCATCCATAGTTGCTGACTCCCGCTGCTAGTAACTACGATACGGAGGGCT 5348
Db 11262 TCGTTCATCCATAGTTGCTGACTCCCGCTGCTAGTAACTACGATACGGAGGGCT 11321
Oy 5349 ACATCTGGCCCCAGTCTGCAATGATACCGGACCCACGCTACCGGCTCCAGATTT 5408
Db 11322 ACATCTGGCCCCAGTCTGCAATGATACCGGACCCACGCTACCGGCTCCAGATTT 11381
Oy 5409 ATCAGCATTAACACGACGACGAGGAGGCGGACGAGAGAGTGGTCTCTGCACTTATC 5468
Db 11382 ATCAGCATTAACACGACGACGAGGAGGCGGACGAGAGAGTGGTCTCTGCACTTATC 11441
Oy 5469 CGCCTCCATCCAGTCTATTAATTTGTCGGGAGGCTAGATAGTATGTCACAGTTAA 5528
Db 11442 CGCCTCCATCCAGTCTATTAATTTGTCGGGAGGCTAGATAGTATGTCACAGTTAA 11501
Oy 5529 TAGTTGGCGACAGTTGTTGCCATTGCTACAGGATGCTGTGACGCTGCTGTTGG 5588
Db 11502 TAGTTGGCGACAGTTGTTGCCATTGCTACAGGATGCTGTGACGCTGCTGTTGG 11561
Oy 5589 TATGGCTTCATTCAGTCTCGGTTGCCAACGATCAAGGAGATTCATGATCCCATGTT 5648
Db 11562 TATGGCTTCATTCAGTCTCGGTTGCCAACGATCAAGGAGATTCATGATCCCATGTT 11621
Oy 5649 GTGCAAAAAAGCGGTAGCTCTTCGCTCCGATCGTTGTCAAGAAATGTTGGCGC 5708
Db 11622 GTGCAAAAAAGCGGTAGCTCTTCGCTCCGATCGTTGTCAAGAAATGTTGGCGC 11681
Oy 5709 AGTGTATCACTCATGTTATGTCAGACACTCATATTTCTTACTGTCAATGCCAT 5768
Db 11682 AGTGTATCACTCATGTTATGTCAGACACTCATATTTCTTACTGTCAATGCCAT 11741
Oy 5769 AAGATGCTTTCTGAGTGTGATGTCATCAACCAAGTCTTGAGAAATGTTGATGCG 5828
Db 11742 AAGATGCTTTCTGAGTGTGATGTCATCAACCAAGTCTTGAGAAATGTTGATGCG 11801
Oy 5829 GCGACCGAGTTGCTCTGGCGCGGCTCAATACGGGATTAATACCGCGCCACATACAGAAC 5888
Db 11802 GCGACCGAGTTGCTCTGGCGCGGCTCAATACGGGATTAATACCGCGCCACATACAGAAC 11861
Oy 5889 TTTTAAAGTGTCTCATTTGAAAAAGCTTCTTGGGGGGAATCTTCAAGATCTTAAC 11921
Db 11862 TTTTAAAGTGTCTCATTTGAAAAAGCTTCTTGGGGGGAATCTTCAAGATCTTAAC 11981
Oy 5949 GCTGTGAGATCAAGTTCGATGTAACCACTCTGTGACACCCAACTGATCTTACGATCTTT 6008
Db 11922 GCTGTGAGATCAAGTTCGATGTAACCACTCTGTGACACCCAACTGATCTTACGATCTTT 11981
Oy 6009 TACTTTACACGAGCTTTCTGGGTGAGCAAAAAACAGAAAGCAAAATGCCGCAAAAAAGG 6068
Db 11982 TACTTTACACGAGCTTTCTGGGTGAGCAAAAAACAGAAAGCAAAATGCCGCAAAAAAGG 12041
Oy 6069 AATAAGGCGACACGGAATGTTAATATCTATCTCTCTTTTCAATATTTATGAG 6128
|||||
```

```
Db 12042 AATAAGGCGACACGAAATGTTGAATACATCTCTCTTTTCAATATTTAG 12101
Oy 6129 CATTTATCAGGGTATTTGTCATGACGCGGATACATATTTGAATATTTAGAAATAA 6188
Db 12102 CATTTATCAGGGTATTTGTCATGACGCGGATACATATTTGAATATTTAGAAATAA 12161
Oy 6189 ACAATAGGGGTTCCGGCACATTTCCCGAAAGTGCAC 6229
Db 12162 ACAATAGGGGTTCCGGCACATTTCCCGAAAGTGCAC 12202

RESULT 2
US-09-943-286-3
; Sequence 3, Application US/09943286
; Patent No. US2002010668A1
; GENERAL INFORMATION:
; APPLICANT: Nunomura, Kiyotada
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
; FILE REFERENCE: GP104-02, JP
; CURRENT APPLICATION NUMBER: US/09/943,286
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 8933
; TYPE: RNA
; ORGANISM: Human Immunodeficiency Virus
; NAME/KEY: source
; FEATURE:
; LOCATION: (1)..(8933)
; OTHER INFORMATION: Sequence of transcripts produced from the BH10
; OTHER INFORMATION: plasmid.
US-09-943-286-3

Query Match 41.7%; Score 2597.8; DB 10; Length 8933;
Best Local Similarity 74.3%; Pred. No. 0;
Matches 1997; Conservative 642; Mismatches 42; Indels 6; Gaps 1;

Oy 1291 CCATGAGAGTGAAGAGAGATATCAGACATTTGTGAGATGGGGTGAATGGGGACCA 1350
Db 5579 CAAGAGAGAGTGAAGAGAGAAUAUACAGACUUGUGAGAUUGGGUGGAGAGUGGGACCA 5638
Oy 1351 TGCTCTTGGGATTTGATGATCTGTAGTGTACAGAAAAATTTGGGTACCGCTTATT 1410
Db 5639 UGCUCUUGGAGUUGUUGAUUGUCUGAGUUGCAGAAAAAUUUGGUCACAGUUAU 5698
Oy 1411 ATGGGTAACCTGTGTGAAGAGCAACCACTCATATTTTGGCATCGATCTTAAG 1470
Db 5699 AUGGGUAACUGUUGAAGAGCAACCACTCATATTTTGGCATCGATCTTAAG 5758
Oy 1471 CATATGATACAGAGTACATTAATGTTGGCCACACATGCTGTACCCACAGACCCA 1530
Db 5759 CAUAGUACAGAGAGUACAUUAUUGUGGCCACACAUUGUCUGUACCCACAGACCCA 5818
Oy 1531 ACCCAAGAGAGTATTTGTTAAATGTGACAGAAAAATTTTACATGTGAAAAATGACA 1590
Db 5819 ACCCAAGAGAGTATTTGTTAAATGTGACAGAAAAATTTTACATGTGAAAAATGACA 5878
Oy 1591 TGGTAGAAGCATGATGAGATTAATGTTTGAAGTGAAGGATTAAGGCTAAGCCATG 1650
Db 5879 UGGUAGAACAGUACAGAGAGAUUAUACUUAUUGGCAAAAGCCUAAAGCCAUUG 5938
Oy 1651 TAAATTAACCCCACTGTGTGTTTAAAGTCACTGATTTGAAATGATGATTAATA 1710
Db 5939 UAAAUUAACCCCACTGTGTGTTTAAAGTCACTGATTTGAAATGATGATTAATA 5998
Oy 1711 CCAATAGTAGTACGGGAGAAATGATTAAGAGAAAGAGATTAATAAACTGCTTTTCA 1770
Db 5999 CCAUUAUAGUAGCGGAGAGAAUAGAUUAGAGAAAGAGAUUAAAAACUGCUCUUA 6058
Oy 1771 ATATCAGCACAAGATTAAGATTAAGTGCAGAAAGAAATATGCTTTTATTAATACTTG 1830
Db 6059 AUUACAGCACAAGCAAGAUUAGAGUUAAGUAGAAAGAAUUAUUGCAUUUUUUUAUUAACUG 6118
|||||
```

OY	2905	AGGCCACACAATTATTTGTCGTGATATATGTCTGCAGCAGCAAGAACAATTTGCTGAGGGCTATTG	2964
Dd	7199	AGGCGCACAAUUAUUUGUCUGGUAUGUUCACAGACAGAACAAAUUUUGUCUGAGGCGUAUUG	7258
OY	2965	AGGGCGAAGCAACTTCTTGTCACACTCCAGAGTCGCGGGCGATCCAAGACAGCTCCAGCGCAAGAA	3024
Dd	7259	AGGCGCAACAGCAUCUUSUCCAACTCCACAGUUCUGGGGCAUCAAAGCAGCUCCACGCAACAA	7318
OY	3025	TCTCGCTGTGGAAAAGATACCTTAAGAGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTG	3084
Dd	7319	UCCUGCGUGGGAAGAGUACCVAAGAUCAACAGCUGCUCGGGGGAUUUGGGGUGUUCUCUG	7378
OY	3085	GAAAACATCATTTGACACACTGCTGCTGCTGGAAATGCTGTGGTAGTAATAAATCTCTG	3144
Dd	7379	GAAAAACCAUUDGACACGUCUGUCUUGGGAUUCGUGGGAGUAUUAUAUUCUCUGG	7438
OY	3145	AACAGATTTGGAAATTAACATGACCTGAGTGGAGGACAGAGAAATTAACAATTACACAA	3204
Dd	7439	AACAGAUUUGGAUUAACAUAGACCUGAGUGGAGGACAGAGAAUUAACAUAUWACAA	7498
OY	3205	GCTTAATACACTCTCTTAATTGAAGAATCCGCAAAACGCAACAAAAAGAAATGACAAAGAT	3264
Dd	7499	GCUTPAACACUCUCUUAUUGAAGAAUCCCAAAACGACGMAAAAAAGAAUGAACAAAGAU	7558
OY	3265	TATTGSAATTAAGATTAATGGGCAAGTTTGGGAATTTGGTAAATTAACATTAACAAATTTGCTGT	3324
Dd	7559	UAUUGGAUUAUGAUAUAUUGGGCAAGUUDUGGAUUGGUUUAACAUACAAAUUDGUCUGU	7618
OY	3325	GGTATATTAATTAATTCATATGATATGATAGTAGAGAGGCTTGGTAGTTTAAGAAATGTTTTG	3384
Dd	7619	GGUAUAUAAAAUUAUUAUUAUUAUUAUUAUGAAGAGCGUUGUAGSUUUUAAGAAUAGUUUUG	7678
OY	3385	CTGTACTTTCTATATGTGAATATAGAGTTAGCGAGGATATTACACATTATCTGTTACAGACC	3444
Dd	7679	CUGAACUUDUCUAGUAUGUAUAGAGUUAAGCAAGGAAUUAACCAUUAUUCGUUCAGACC	7738
OY	3445	ACCTCCCAATCCCGAGGGGAGCCCGACAGGCCCCGAAGGAATTAAGAAGAGATGGAGAGA	3504
Dd	7739	ACUCCCAAUCCGAGGGGACCCGACAGCGCCGCAAGAAUUAAGAAGAGUGGAGAGA	7798
OY	3505	GAGACAGAGACAGATCCATTGCTATATGTAGAGGATCTTACGACTTATCTGGAGCATC	3564
Dd	7799	GAGACAGAGACAGAUCCAUUCGAUUAUGUUAACGAGACUUAACUUAUUCUGGAGCAUUC	7858
OY	3565	TGCGGAGCGTGTGCTCTTACGATACCAOCGCTTGAGACATTAACCTTGATTTGAAGA	3624
Dd	7859	UGCGGACCCUGUGCCUUCUAGUCUACCCAGCGCUGUGAGACAUACUUAUUGAUUGAAGCA	7918
OY	3625	GGATTGTGGAATCTCTGGAGCGCAGGGGGTGGGAACCCCTCAAAATATTGGTGGAAATCCC	3684
Dd	7919	GGAUUUGGAACUUCUGGACGCGAGGGGGUGSSAACCCUCAAUUUGUGGGAUVCUC	7978
OY	3685	TACAGTATTGGAGTCAAGAACTTAAGATATAGTCTGTATTAATTGCTCAATGCCCACGCA	3744
Dd	7979	UACAGUUAUUGAGACGACGAAUAAAUUAUGUCUUAUGUCUCCAUUAGCCACAGCUA	8038
OY	3745	TACAGATAGCTGAGGGGACAGATAGSGTTATACAACTATTACAAGCAGCTTATACAGCTA	3804
Dd	8039	UAGCAGUAGCUGAGGGGACAGUAGGGUUAUACAAUACUACAAAGAGCUCUUAUACGUA	8098
OY	3805	TTGCCCACTACCTAGAGAAATTAAGACAGGGCTTGGAAAGATTTGGCTATTAATGAGTGGT	3864
Dd	8099	UUGCCCAUAUACCUAGAGAAUUAUAGACAGGGCUCUGSAAAGAUUUUUGCUUAUUAAGUUGGU	8158
OY	3865	GGCAAGTGTGTCAAAAAGTACTGTGATTGGATGTGGCTCTGTAAAGGAAAAAGAAATGAGACA	3924
Dd	8159	GGCAAGUGGUCAAAAAGUAGUUGUGUGUGGUCUGCUUAAGGAAAAAGAAUAGAGACA	8218
OY	3925	GCTAGGCCACACAGATGTGGGTGGGACGACGATATCTCGAGATCTAGA	3971
Dd	8219	GCUBAGCCACACACAGUUGGGGUGGAGAGCAUUCUCCAGGACCUAGA	8265

QY	2123	ATGCTAAACCATTAATAGTACAGCTGAACACACTCTGTAGAAATTAATTGTACAAGACCCA	2184
Db	6419	AUGCCAAAACCCUUAUACUAGCGCAACCAUUCGUAGAAAUUAUUGUCAAAGACCCA	6478
QY	2185	ACAACAAATACAGAAAAGTATCCGTATCCAGAGGGGACCCAGGGGAGACATTTGTTCCAA	2244
Db	6479	ACAACAAATACAGAAAAGTATCCGTATCCAGAGGGGACCCAGGGGAGACATTTGTTCCAA	6538
QY	2245	TAGAAAAAATAGAAATATGAGACAAAGACATTTGTACATTAGTAGACAAAATGGAATG	2304
Db	6539	UAGGAAAAAUUGGAAAUUUGAGACAAAGCACAUUGUACAUAUAGUAGACAAAUGGAAUA	6598
QY	2305	CCACCTTTAAACAGATAGCTAGTGCAGCAATTAAGAGAAACATTTGGAAATTAATTAACAATAA	2364
Db	6599	ACACUUAACACAGUACACUADUAGCAAAUUAAGAAACAUAUUGGAAUAUUAACAAACAA	6658
QY	2365	TCTTTTAAGCAATCCTCAGAGGGGACCCAGAAATTTGTAAACGACAGTTTAATTGTGGAG	2424
Db	6659	UCUUUAACAGACGCTCCAGAGGGGACCCAGAAUUGUUAACGACAGUUUUUAUUGUGGAG	6718
QY	2425	GGGAAATTTTCTACGTATATTCACACAACTGTTTAATAGTACTTGTTTAATAGTACTT	2484
Db	6719	GGGAUUUUUUUCACGDUAAUUCACACACACGUUUUAUAGUACUUGUUUAUUGUACUU	6778
QY	2485	GGAGTACTGAAGGGTCAAAATAACACTGAAGGAAGTGCACACATCAACATCTCCATGCAGAA	2544
Db	6779	GGAGUACUAAAAGGCTUCAAUAUACUGAAAGGAGUACACAUCACCUCUCCAGUCAGAA	6838
QY	2545	TAAACAAATTTATTAACATGTGGCGAGGAAGTATAGGAAAAAGCAATGTATGCCCTCCATCA	2604
Db	6839	UAAAACAAAUUUAACAUUGUGGCGAGGAAGGAGAAAGCAUAGUATUCCUCCUCANCA	6898
QY	2605	GTCGACAAATTAAGATGTCATCAATTAATTAAGTGGGCTCTTAATTAACAAGATGATGGTGA	2664
Db	6899	GUGGACAAAUUAGAUUGUUCACUCAAUAUUAACAGGCGUCUUAUUAACAAGAAUUGUGUA	6958
QY	2665	ATAACACAATGGGCTCCGAGATCTTTCAGACCTGGAGAGGCGGATATGAGGACAATTGGA	2724
Db	6959	AUAGCAACAUAUGAGUCCGAGAUUCUACAACCCUGGAGGAGAGUAUUGAGGACAAUUGGA	7018
QY	2725	GAAAGTAAATTTATTAATATTAAGTATGTAATAATTGAACCTTGTAGGATAGCACCACCA	2784
Db	7019	GAAAGTAAATTTATTAATATTAAGTATGTAATAATTGAACCTTGTAGGATAGCACCACCA	7078
QY	2785	AGGCAAGAAGAGAGTGGTGCAGAGAGAAAAAAGACAGTGGGAATAGAGACTTTGTTC	2844
Db	7079	AGGCAAGAAGAGAGTGGTGCAGAGAGAAAAAAGACAGTGGGAATAGAGACTTTGTTC	7138
QY	2845	TTGGGTTCTTGGGAGCAGCAGAAAGCACTATAGGGCTGCACCTCAATACGCTGACGGTAC	2904
Db	7139	UUGGGUUCUUGGGAGCAGCAGAAAGCACUUAUAGGCGCGCCUCCAUGACGUGAGGUA	7198
QY	2905	AGGCGAGCAATTTATTTGCTATGTATAGTACGACAGAGAAACAATTTGGTGGGGCTATTG	2964
Db	7199	AGGCGAGCAATTTATTTGCTATGTATAGTACGACAGAGAAACAATTTGGTGGGGCTATTG	7258
QY	2965	AGGCGCAACAGCATCTGTTGCACTCACAGTCTGGGGCATCAAAACGCTCCAGGCAAGAA	3024
Db	7259	AGGCGCAACAGCATCTGTTGCACTCACAGTCTGGGGCATCAAAACGCTCCAGGCAAGAA	7318
QY	3025	TCTGCGCTGTGGAAAGATTAAGTAAAGATCAACAGCTCTCTGGGATTTGGGGTGTGCTG	3084
Db	7319	UCCUGGCGUGGAAAAGAUACCUAAAAGUACAACAGCUCUUGGGGAUUUGGGUGGCGUC	7378
QY	3085	GAAGAACTATTATTCACACTGCTGTGCTTGGATGTAGTGTGGAGTAAATTAATCTGTG	3144
Db	7379	GAAGAACTATTATTCACACTGCTGTGCTTGGATGTAGTGTGGAGTAAATTAATCTGTG	7438
QY	3145	AACAGATTTGGAATTAACATGACCTGATGAGTGGGACAGAGAAATTAACATTAACAA	3204
Db	7439	AACAGATTTGGAATTAACATGACCTGATGAGTGGGACAGAGAAATTAACATTAACAA	7498

QY	3205	GCTTAATACACTCCTTATTGAGATCCGAATAAACGACAGAATAAAGATGAACAAGATT	3264
Db	7499	GCUTAAUACACUCUUUAUUUGAAGAAACCGCAAAACCGCACGAAGAAAAGAUACAAGAADU	7558
QY	3265	TATTGGAAATTGTAAATTTGGGCAAGTTTGAGAAATTGGTTTAATACATAAATTTGGCTGT	3324
Db	7559	UADUGGAUUDUGAATAAUVGGGCCAAGUUDUGGGAUUDUGUUAACATAACAAUUDGGCUUD	7618
QY	3325	GGTATATAAAAATTTATCATATGATAGTAGAGGCTTGGTAGGTTTAAAGATTAGTTTTTG	3384
Db	7619	GGUADVAAAAUUUVAUUCUAUAUAGUAGAGAGCUUGUGAGUUYUUAAGAUAUAGUUUUG	7678
QY	3385	CTGTACTTTCATATGTGAATAGAGTTTGGCAGGGGATTTACCATTATGCTTTCAGACCC	3444
Db	7679	CUGACACUUUCUGUAGUUAAGAUUDUGGACAGGGAUUDUCCACUAUUDUGUUUCACACCCC	7738
QY	3445	ACCTCCCAATCCCGAGGGGACCCGACAGGCCCCGAAGGAATAGAAAGAAAGGTGAGAGA	3504
Db	7739	ACUCCCAAUUCCGAGGGGACCCGACAGGCCCGCAAGGAUAUAGAAAGAAAGUGGAGANA	7798
QY	3505	GAGACAGAGACAGATCCATTTCGATTAGTGAACGATTCCTTAGCATTATCTGGGAGATC	3564
Db	7799	GAGACAGAGACAGAUCCAUUCGAUUDAGUAAACGAGAUCCUAGCAGCAUUAUCUGGACGAC	7858
QY	3565	TGCGAGAGCTGTGCCTCTTCAGCTACACCGCTTGAGAGACTTACTCTTGATTGTAAAGA	3624
Db	7859	UGCAGAGCCUUBGCUUCUCAACUACUACCCGCUUGAGACAGCAUUCUUGAUUDGUAACGA	7918
QY	3625	GGATTGTGGAATCTTGTGGAGCGCAGGGGGTGGGAAGCCCTCAAATATTGTGGAAATCTCC	3684
Db	7919	GGAUUDUGGAACUUUCUGGAGACGACAGGGGGUGGGGAAGCCCCCAAAUUAUUDUGUGAAUUCUC	7978
QY	3685	TACAGTATTGGAGTCAGGAACCTAAAGAAATAGTCTGTTAACCTTGCTCAATGCCACAGCCA	3744
Db	7979	UACAGUUDUGGAGAGCAGACAGCUAAAGAUUDGUCUUDUGGCUCAAUUDGCCACACGCTA	8038
QY	3745	TAGAGATAGCTGGAAGGAGACAGATAGGGTTTAAAGATTAACAGCAAGCTTATAGAGCTA	3804
Db	8039	UAGCAGATAGCUGAGGGGAGCAGAUAGGSUUUAAGAUAGUACAGGAGCUUUAUAGACCTA	8098
QY	3805	TTTCGCCACATACCTAGAAATTAAGACAGGGCTTGGAAAAGATTTTTGCTATAAGATGGGT	3864
Db	8099	TUCCCCACAUAUCCAGAAAGAAUUAAGACAGGCGUUGGAAAAGAUUUDGCUAUAAGUAGGU	8158
QY	3865	GGCAAGGTGTCAAAACCTACTGTGATGTGATGGCCCTGTGAAGGAAGAATGAGAGCA	3924
Db	8159	GGCAAGGTGTCAAAACCTACTGTGATGTGATGGCCCTGTGAAGGAAGAATGAGAGCA	8218
QY	3925	GCTAGGCGACAGCAGAGATGGGTGGGAGCAGTATCTCGAANTGAGA	3971
Db	8219	GCUTAGGCGACAGCAGAGATGGGTGGGAGCAGTATCTCGAANTGAGA	8265
RESULT 5			
US-09-737-190A-1			
; Sequence 1, Application US/09737190A			
; Patent No. US20020102545A1			
GENERAL INFORMATION:			
APPLICANT: Shinya, Tetsuo			
TITLE OF INVENTION: A Method for Changing a Target Array, a Method for Analyzing			
TITLE OF INVENTION: a Structure, and an Apparatus, a Storage Medium and a			
FILE REFERENCE: JP9199902700S1 (14043)			
CURRENT APPLICATION NUMBER: US/09/737,190A			
NUMBER OF SEQ ID NOS: 2			
SOFTWARE: Patentin Ver. 2.1			
SEQ ID NO 1			
LENGTH: 9719			
TYPE: DNA			
ORGANISM: Human Immunodeficiency virus type 1			
US-09-737-190A-1			

Query Match		41.6%;	Score 2593;	DB 10;	Length 9719;
Best Local Similarity		98.1%;	Pred. No. 0;		
Matches 2636;		Conservative	0;	Mismatches	45;
				Indels	6;
				Gaps	1;
QY	1291	CCATGAGAGTGAAGAGAAATGATCAGCAGCTTGTGGAGATGGGGGTGGAATGGGGACCA	1350		
DB	6223	CAATGAGAGTGAAGAGAAATGATCAGCAGCTTGTGGAGATGGGGGTGGAATGGGGACCA	6282		
QY	1351	TGCTCCCTGGGATATGTGATCTGTAGTGTACAGAAATTTGGGGTCCACCGCTATT	1410		
DB	6283	TGCTCCCTGGGATATGTGATCTGTAGTGTACAGAAATTTGGGGTCCACCGCTATT	6342		
QY	1411	ATGGGGTACCTGTGTGAAGAGCAACACACCTCTATTGTTGTGCATCAGATGCTAAG	1470		
DB	6343	ATGGGGTACCTGTGTGAAGAGCAACACACCTCTATTGTTGTGCATCAGATGCTAAG	6402		
QY	1471	CATATGATACAGAGGTACATTAATGTTGGGCCACATGCTGTGTACCCACAGACCCA	1530		
DB	6403	CATATGATACAGAGGTACATTAATGTTGGGCCACATGCTGTGTACCCACAGACCCA	6462		
QY	1531	ACCCACAGAGTATGTATGTAATGTGACAGAAATTTTAACTGTGGAAATATGACA	1590		
DB	6463	ACCCACAGAGTATGTATGTAATGTGACAGAAATTTTAACTGTGGAAATATGACA	6522		
QY	1591	TGTTAGACAGATGATGAGATATATCAGTTTATGGGATCAAAAGCCATAAGCATGTG	1650		
DB	6523	TGTTAGACAGATGATGAGATATATCAGTTTATGGGATCAAAAGCCATAAGCATGTG	6582		
QY	1651	TAAATTTAACCCCACTGTGTAGTTTAAAGTGCATGATTTGAAGATGATCTAATA	1710		
DB	6583	TAAATTTAACCCCACTGTGTAGTTTAAAGTGCATGATTTGAAGATGATCTAATA	6642		
QY	1711	CCATAGTATGAGGGGAGATGATTAATGAGAAAGAGATATAAAACGCTCTTTCA	1770		
DB	6643	CCATAGTATGAGGGGAGATGATTAATGAGAAAGAGATATAAAACGCTCTTTCA	6702		
QY	1771	ATATCAGCACAAGCATAAGATTAAGTGCAGAAAGATATGATCTTTTATAAATCG	1830		
DB	6703	ATATCAGCACAAGCATAAGATTAAGTGCAGAAAGATATGATCTTTTATAAATCG	6762		
QY	1831	ATATAGTATCCAAATAGATA-----TACCAGCTATAGTTGATTAAGTGAACACCTGAG	1884		
DB	6763	ATATAGTATCCAAATAGATA-----TACCAGCTATAGTTGATTAAGTGAACACCTGAG	6822		
QY	1885	TCATTACACAGGCGCTGCCAAGGATACCTTTGACCCCAATCCCAATATATTTGGCC	1944		
DB	6823	TCATTACACAGGCGCTGCCAAGGATACCTTTGACCCCAATCCCAATATATTTGGCC	6882		
QY	1945	CGGCTGTTTTGCAATCTTAAATGTATTAATTAAGACGTTCAATGGAACAGACCATGTA	2004		
DB	6883	CGGCTGTTTTGCAATCTTAAATGTATTAATTAAGACGTTCAATGGAACAGACCATGTA	6942		
QY	2005	CAAAATGTCAGCAGATACATATGTACACATGGAATTCAGGCCAGTATATCAACTAATGC	2064		
DB	6943	CAAAATGTCAGCAGATACATATGTACACATGGAATTCAGGCCAGTATATCAACTAATGC	7002		
QY	2065	TGTTAAATGGAGCTACGACAAGAGATGTATTAATTAAGATTCCTCCCAATTCACAGACA	2124		
DB	7003	TGTTAAATGGAGCTACGACAAGAGATGTATTAATTAAGATTCCTCCCAATTCACAGACA	7062		
QY	2125	ATGCTAAACCATATATATAGAGTGAACATATCTGTAGAAATTAATGTACAAGACCA	2184		
DB	7063	ATGCTAAACCATATATATAGAGTGAACATATCTGTAGAAATTAATGTACAAGACCA	7122		
QY	2185	ACAACATACAGAAAAAGTATCCGTATCCAGAGGGGACCAAGGAGACATTTGTTACAA	2244		
DB	7123	ACAACATACAGAAAAAGTATCCGTATCCAGAGGGGACCAAGGAGACATTTGTTACAA	7182		
QY	2245	TAGGAAAAATGGAATATGAGACAAGACATTTGTAATATAGTAGACAAAAATGGAATG	2304		
DB	7183	TAGGAAAAATGGAATATGAGACAAGACATTTGTAATATAGTAGACAAAAATGGAATG	7242		
QY	2305	CCACTTTAAAAACAGATAGCTAGCAAAATTAAGAGAACAAATTTGAAATTAATTAACAATA	2364		

DB	7243	ACACTTTAAACAGATAGCTAGCAAAATTAAGAGAACAAATTTGGAATATATAAACAAATA	7302		
QY	2365	TCTTTAAGCAATCTCAGAGAGGGACCCAGAAATTTGAACGACAGTTTAAATGTGAG	2424		
DB	7303	TCTTTAAGCAATCTCAGAGAGGGACCCAGAAATTTGAACGACAGTTTAAATGTGAG	7362		
QY	2425	GGGAATTTTCTAGCTATATTCACACAAGTGTAAATAGTACTGTTTAAATATACTACT	2484		
DB	7363	GGGAATTTTCTAGCTATATTCACACAAGTGTAAATAGTACTGTTTAAATATACTACT	7422		
QY	2485	GGAGTACGAAGGGTCAAAATTAACACTGAAGAAAGTGAACACATCAGCTCCATGAGAA	2544		
DB	7423	GGAGTACGAAGGGTCAAAATTAACACTGAAGAAAGTGAACACATCAGCTCCATGAGAA	7482		
QY	2545	TAAACCAATTTAATAACATGTGGCAGAGATGAGAAAGCAATGTATCCCTCCATCA	2604		
DB	7483	TAAACCAATTTAATAACATGTGGCAGAGATGAGAAAGCAATGTATCCCTCCATCA	7542		
QY	2605	GTGGCAATTAAGATGTTCAATCAATATTTACTGGGCTGCTATTACAGAGATGGTGTA	2664		
DB	7543	GTGGCAATTAAGATGTTCAATCAATATTTACTGGGCTGCTATTACAGAGATGGTGTA	7602		
QY	2665	ATAACCAATGGGTCCGAGATCTTCAGACCTGGAGAGGGCATATGAGGCAATTTGGA	2724		
DB	7603	ATAACCAATGGGTCCGAGATCTTCAGACCTGGAGAGGGCATATGAGGCAATTTGGA	7662		
QY	2725	GAAGTAATTTATTAATATTAAGATGTAATAAATTTGAACATTTAGAGTACCCACCA	2784		
DB	7663	GAAGTAATTTATTAATATTAAGATGTAATAAATTTGAACATTTAGAGTACCCACCA	7722		
QY	2785	AGGCCAAGAGAAATGTGTGCAGAGAGAAAAAGCAGTGGGAATAGAGCTTTGTCC	2844		
DB	7723	AGGCCAAGAGAAATGTGTGCAGAGAGAAAAAGCAGTGGGAATAGAGCTTTGTCC	7782		
QY	2845	TTGGCTTTTGGGACACAGAGAACACTATGCGCTGCACGTCATGACGCTGACGATAC	2904		
DB	7783	TTGGCTTTTGGGACACAGAGAACACTATGCGCTGCACGTCATGACGCTGACGATAC	7842		
QY	2905	AGGCCAGCAATTTATTTCTGATATATGTGCAGACAGACAGAACATTTGCTGAGGCTATTG	2964		
DB	7843	AGGCCAGCAATTTATTTCTGATATATGTGCAGACAGACAGAACATTTGCTGAGGCTATTG	7902		
QY	2965	AGGGCAACAGATCTGTGCAACACAGTCTGGGGGATCAACAGCTCCAGGCAAGAA	3024		
DB	7903	AGGGCAACAGATCTGTGCAACACAGTCTGGGGGATCAACAGCTCCAGGCAAGAA	7962		
QY	3025	TCTGGGCTGTGGAAGATACCTAAAGATCAACACGCTCTGGGGAATTTGGGTTGCTCG	3084		
DB	7963	TCTGGGCTGTGGAAGATACCTAAAGATCAACACGCTCTGGGGAATTTGGGTTGCTCG	8022		
QY	3085	GAANAATCAATTTGCACCACTGCTGCTTGGGAATGTAGTTGGATTAATAATCTGCG	3144		
DB	8023	GAANAATCAATTTGCACCACTGCTGCTTGGGAATGTAGTTGGATTAATAATCTGCG	8082		
QY	3145	AACGATTTGGGAATTAACATGACCTGAGTGGAGTGGGACAGAGAAATTAACAATTACACA	3204		
DB	8083	AACGATTTGGGAATTAACATGACCTGAGTGGAGTGGGACAGAGAAATTAACAATTACACA	8142		
QY	3205	GCTTAATACACTCTTAATTAAGAAATGCAAAAACACAGCAAGAAAAGATGAACAAGAT	3264		
DB	8143	GCTTAATACACTCTTAATTAAGAAATGCAAAAACACAGCAAGAAAAGATGAACAAGAT	8202		
QY	3265	TATTTGAATTAAGATTAATGGCAAGTTTGTGAATTTGTTTAACATAACAAATTTGGCTGT	3324		
DB	8203	TATTTGAATTAAGATTAATGGCAAGTTTGTGAATTTGTTTAACATAACAAATTTGGCTGT	8262		
QY	3325	GGTATATTAATTAATTAATGAATAGTAGAGGCTGTGGTATAGTAATAGTATTG	3384		
DB	8263	GGTATATTAATTAATTAATGAATAGTAGAGGCTGTGGTATAGTAATAGTATTG	8322		
QY	3385	CTGTACTTTCTATAGTGAATAGAGTTAGCAGAGGATATTCACCAATTAATCTTTGACAGCC	3444		

Db 8323 CTGTACTTCTGTATAGTGAATAGAGTATAGGAGGGATATTCACATTATCGTTTCAGACCC 8382
Qy 3445 ACCTCCCAATCCCGAGGGGAGCCGACAGGCCCCGAAGAAATAGAAAGAGGTGAGAGA 3504
Db 8383 ACCTCCCAACCCCGAGGGGAGCCGACAGGCCCCGAAGAAATAGAAAGAGGTGAGAGA 8442
Qy 3505 GAGACAGAGACAGATTCATTCGATTAAGTGAACGGATCCTTAGCACTTATCTGGGACATC 3564
Db 8443 GAGACAGAGACAGATTCATTCGATTAAGTGAACGGATCCTTAGCACTTATCTGGGACATC 8502
Qy 3565 TGGCGAGCCCTGTGCTCTTACAGTACACCGCTTGAGAGACTTACTCTGATTTGAACA 3624
Db 8503 TGGCGAGCCCTGTGCTCTTACAGTACACCGCTTGAGAGACTTACTCTGATTTGAACA 8562
Qy 3625 GGATTTGTGAACCTCTGAGGAGCGAGGGGGTGGGAAGCCCTCAATATTTGGTGAATCTCC 3684
Db 8563 GGATTTGTGAACCTCTGAGGAGCGAGGGGGTGGGAAGCCCTCAATATTTGGTGAATCTCC 8622
Qy 3685 TACAGTATGTGAGTACAGAACTAAAGAAATAGTGTCTGTTAACTTCTCAATCCACAGCCA 3744
Db 8623 TACAGTATGTGAGTACAGAACTAAAGAAATAGTGTCTGTTAACTTCTCAATCCACAGCCA 8682
Qy 3745 TAGCAGTAGCTGAGGGGAGCAGATAGGGTTATAGAAGTATTACAGCAGCTTATAGAGCTA 3804
Db 8683 TAGCAGTAGCTGAGGGGAGCAGATAGGGTTATAGAAGTATTACAGCAGCTTATAGAGCTA 8742
Qy 3805 TTGCGCACATACCTAGAAAGATAAGACAGAGGCTTGGAAGGATTTGCTATATAGATGGGT 3864
Db 8743 TTGCGCACATACCTAGAAAGATAAGACAGAGGCTTGGAAGGATTTGCTATATAGATGGGT 8802
Qy 3865 GGCAAGTGTCTCAAAAAGTATAGTGTGATGGCTCTGTAAGGGAAGAAATAGAGAGA 3924
Db 8803 GGCAAGTGTCTCAAAAAGTATAGTGTGATGGCTCTGTAAGGGAAGAAATAGAGAGA 8862
Qy 3925 GCTGAGCAGCAGCAGATAGGGGTGGAGCAGATCTGAGATCTAGA 3971
Db 8863 GCTGAGCAGCAGCAGATAGGGGTGGAGCAGATCTGAGATCTAGA 8909

RESULT 6
US-10-000-511A-1
: sequence 1, Application US/10000511A
: Publication No. US20030158131A1
: GENERAL INFORMATION:
: APPLICANT: Aldovini, Anna
: TITLE OF INVENTION: DNA VECTORS CONTAINING MUTATED HIV PROVRUSES
: FILE REFERENCE: 101353-165
: CURRENT APPLICATION NUMBER: US/10/000, 511A
: CURRENT FILING DATE: 2001-10-23
: PRIOR APPLICATION NUMBER: 60/242, 589
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/243, 432
: PRIOR FILING DATE: 2000-11-28
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 9719
: TYPE: DNA
: ORGANISM: Human immunodeficiency virus
US-10-000-511A-1

Query Match 41.68; Score 2593.; DB 12; Length 9719;
Best Local Similarity 98.18; Pred. No. 0;
Matches 2636; Conservative 0; Mismatches 45; Indels 6; Gaps 1;

Qy 1291 CCATGAGAGTGAAGAGAGATATACACATTTGTGAATAGGGGTGGAATAGGGGACCA 1350
Db 6223 CAATGAGAGTGAAGAGAGATATACACATTTGTGAATAGGGGTGGAATAGGGGACCA 6282
Qy 1351 TGCCTCTGGATATGATGATCTGTAGTGTACAGAAAAATTTGGGTCACTCATTT 1410
Db 6283 TGCCTCTGGATATGATGATCTGTAGTGTACAGAAAAATTTGGGTCACTCATTT 6342

Qy 1411 ATGGGGTACCTGTGTGGAAGAACCAACCACTCATTTTGTGATCAGATGTAAAG 1470
Db 6343 ATGGGGTACCTGTGTGGAAGAACCAACCACTCATTTTGTGATCAGATGTAAAG 6402
Qy 1471 CATATGATACAGAGTACATATATTTTGGGCAACACATGGCTGTGTACCCAGACCCCA 1530
Db 6403 CATATGATACAGAGTACATATATTTTGGGCAACACATGGCTGTGTACCCAGACCCCA 6462
Qy 1531 ACCCACGAAGATGTATTTGTAATGTGACAGAAATTTTAAACATGTGAAAAATGACA 1590
Db 6463 ACCCACGAAGATGTATTTGTAATGTGACAGAAATTTTAAACATGTGAAAAATGACA 6522
Qy 1591 TGTAGAACAGATGATGAGATTAATCAAGTTTATGGATCAAAAGCTTAAGCCATGTG 1650
Db 6523 TGTAGAACAGATGATGAGATTAATCAAGTTTATGGATCAAAAGCTTAAGCCATGTG 6582
Qy 1651 TAAATTAACCCCACTCTGTGTGTGTTTAAAGTCACTGTATTTGAAGAAATGATCTATA 1710
Db 6583 TAAATTAACCCCACTCTGTGTGTGTTTAAAGTCACTGTATTTGAAGAAATGATCTATA 6642
Qy 1711 CCAATAGTAGGCGGAGAGATGAATTAAGGAAAGAGAGATTAATAAAGCTGCTTTCA 1770
Db 6643 CCAATAGTAGAGCGGGAGAGATGAATTAAGGAAAGAGAGATTAATAAAGCTGCTTTCA 6702
Qy 1771 ATATCAGACAGACATTAAGATTAAGTGTGACAGAAAGATATGCTTTTATTAACCTG 1830
Db 6703 ATATCAGACAGACATTAAGATTAAGTGTGACAGAAAGATATGCTTTTATTAACCTG 6762
Qy 1831 ATATAGTACCAATATGATTA-----TACACACTTATGATTAAGTGTGACACCTGAG 1884
Db 6763 ATATATATACCAATATGATTAAGTGTGACACCTTATTAAGTGTGACACCTGAG 6822
Qy 1885 TCATTAACAGAGCCTGTCCAAAGGATCTTTGAGCCAAATTCCTATTAATTTGTGCC 1944
Db 6823 TCATTAACAGAGCCTGTCCAAAGGATCTTTGAGCCAAATTCCTATTAATTTGTGCC 6882
Qy 1945 CGGCTGTTTGGCATTTCTAAATATTAATTAAGACCTTCAATGGAACGACCATGTA 2004
Db 6883 CGGCTGTTTGGCATTTCTAAATATTAATTAAGACCTTCAATGGAACGACCATGTA 6942
Qy 2005 CAAATGTACAGCAGTCAATGTATACATGGAATCAGGCCAGTGTATCAATCACTACCTGC 2064
Db 6943 CAAATGTACAGCAGTCAATGTATACATGGAATCAGGCCAGTGTATCAATCACTACCTGC 7002
Qy 2065 TGTAAATGGCAGTCTAGCAGAAAGAGATGTAATTAAGTGTGCAATTTCCAGACA 2124
Db 7003 TGTAAATGGCAGTCTAGCAGAAAGAGATGTAATTAAGTGTGCAATTTCCAGACA 7062
Qy 2125 ATGCTAAACCATATATATGACAGTGAACACATCTGTGAATTAATTTGTAAGACCCA 2184
Db 7063 ATGCTAAACCATATATATGACAGTGAACACATCTGTGAATTAATTTGTAAGACCCA 7122
Qy 2185 ACAACATPACAGAAAAAGTATCCGTATCCAGAGGGAGCAAGGAGAGCATTTGTTCAA 2244
Db 7123 ACAACATPACAGAAAAAGTATCCGTATCCAGAGAGACCAAGGAGAGCATTTGTTCAA 7182
Qy 2245 TAGAAAAATAGAAAAATATGAGACAGACATTTGTAATTAAGTGAAGAAATGAGATG 2304
Db 7183 TAGAAAAATAGAAAAATATGAGACAGACATTTGTAATTAAGTGAAGAAATGAGATG 7242
Qy 2305 CCATTTAAACAGATATGCTAGCAATTAAGAGACAAATTTGGAATTAATTAACATAA 2364
Db 7243 ACATTTAAACAGATATGCTAGCAATTAAGAGACAAATTTGGAATTAATTAACATAA 7302
Qy 2365 TCTTTAAGCAATCCTCGAGGAGGAGCCAGAAATTTGACGACAGTTTAATTTGAGAG 2424
Db 7303 TCTTTAAGCAATCCTCGAGGAGGAGCCAGAAATTTGACGACAGTTTAATTTGAGAG 7362
Qy 2425 GGGAAATTTTCTACTGTATTAATTAACACAGCTTTTAATTAATTAATTAATTAATTA 2484
Db 7363 GGGAAATTTTCTACTGTATTAATTAACACAGCTTTTAATTAATTAATTAATTAATTA 7422
Qy 2485 GGAAGTACTGAAGGGTCAATTAACACTGAAGAAAGTGAACAAATCCACTCCATGAGAA 2544

Db	7423	GGAGTACTGAAGGGTCAAAATTAACACTGAAGGAAGTGAACCAATACCTCCCATGACGAA	7482
QY	2545	TAAACAATTTATTAACATGTGGCAGAGATAGGAAAGCAATGTATGCCCCCTCCATCA	2604
Db	7483	TAAACAATTTATTAACATGTGGCAGAGAAAGTAGGAAAAAGCAATGTATGCCCCCTCCATCA	7542
QY	2605	GTGGACAATTTAGATGTTTCATCAAAATATTACTGGGCTGCTATTAAACAAGAGATGGTGT	2664
Db	7543	GTGGACAATTTAGATGTTTCATCAAAATATTACTGGGCTGCTATTAAACAAGAGATGGTGT	7602
QY	2665	ATTAACAACATTTGGTCCGAGATCTTTCAGACCTGGAGAGGCGATATGAGGACAATTGGA	2724
Db	7603	ATTAACAACATTTGGTCCGAGATCTTTCAGACCTGGAGAGGCGATATGAGGACAATTGGA	7662
QY	2725	GAACTGAATTTTATTAATTTATTAATTAAGTAAATTTGAACCTTATGAGATGACACCCACCA	2784
Db	7663	GAACTGAATTTTATTAATTTATTAATTAAGTAAATTTGAACCTTATGAGATGACACCCACCA	7722
QY	2785	AGGCAAAAGAGAAGTGGTGCAGAGAGAAAAAGAGCAGTGGGAATAGAGACTTTGTTTCC	2844
Db	7723	AGGCAAAAGAGAAGTGGTGCAGAGAGAAAAAGAGCAGTGGGAATAGAGACTTTGTTTCC	7782
QY	2845	TTGGGTTCTTGGGAGCAGCAGGAAGCAGTATGGGCTGCAGCTCATATGACCTGACGGTAC	2904
Db	7783	TTGGGTTCTTGGGAGCAGCAGGAAGCAGTATGGGCTGCAGCTCATATGACCTGACGGTAC	7842
QY	2905	AGGCGACAGCAATTTATGTCGTATTTAGTGGCAGCAGCAAGCAATTTGCTGAGGCTATTG	2964
Db	7843	AGGCGACAGCAATTTATGTCGTATTTAGTGGCAGCAGCAAGCAATTTGCTGAGGCTATTG	7902
QY	2965	AGGCGCAACAGCATCTGTGCACTCAGCTGTGGGCGATCAAAACAGCTCCAGGCAAGAA	3024
Db	7903	AGGCGCAACAGCATCTGTGCACTCAGCTGTGGGCGATCAAAACAGCTCCAGGCAAGAA	7962
QY	3025	TTCCTGGCTGTGTGAAGATACCTAAAGGATCAACAGCTCTGTGGGATTTGGGCTTGTCTG	3084
Db	7963	TTCCTGGCTGTGTGAAGATACCTAAAGGATCAACAGCTCTGTGGGATTTGGGCTTGTCTG	8022
QY	3085	GAAACACTCATTTTGACACAGCTGTGGCTTGGGATGGTAGTTGGAGTAAATTAATCTCTGG	3144
Db	8023	GAAACACTCATTTTGACACAGCTGTGGCTTGGGATGGTAGTTGGAGTAAATTAATCTCTGG	8082
QY	3145	AACAGATTTGGAAATTAACATGACGCTGGATGGAGTGGGACAGAGAAATTAACAAATTAACAA	3204
Db	8083	AACAGATTTGGAAATTAACATGACGCTGGATGGAGTGGGACAGAGAAATTAACAAATTAACAA	8142
QY	3205	GCTTAATACACTCTTTAATTAAGATGCGCAAAACCGCAAGAAAGATGACAGAT	3264
Db	8143	GCTTAATACACTCTTTAATTAAGATGCGCAAAACCGCAAGAAAGATGACAGAT	8202
QY	3265	TATTGGAATTAAGATTAATGGGCAAGTTTGGGATTTGGTTTAACATTAACAAATTTGGCTGT	3324
Db	8203	TATTGGAATTAAGATTAATGGGCAAGTTTGGGATTTGGTTTAACATTAACAAATTTGGCTGT	8262
QY	3325	GGTATATTAATTTATTCATATGATGATAGAGGCTTGTAGTGAAGTTAAGATATATTTTGG	3384
Db	8263	GGTATATTAATTTATTCATATGATGATAGAGGCTTGTAGTGAAGTTAAGATATATTTTGG	8322
QY	3385	CTGTACTTTTCTATAGTGAATAGAGTTAGGCAAGGATTTACCACTTATCGTTTACAGACC	3444
Db	8323	CTGTACTTTTCTATAGTGAATAGAGTTAGGCAAGGATTTACCACTTATCGTTTACAGACC	8382
QY	3445	ACCTCCCAATCCCGAGGGGACCCGACAGGCCCGCAAGGAATTAAGAAAGAGTGGACAGA	3504
Db	8383	ACCTCCCAATCCCGAGGGGACCCGACAGGCCCGCAAGGAATTAAGAAAGAGTGGACAGA	8442
QY	3505	GAGACAGAGACAGATTCATTCGATTAAGTGAAGGATCCTTAGCACTTATCTGGGACATC	3564
Db	8443	GAGACAGAGACAGATTCATTCGATTAAGTGAAGGATCCTTAGCACTTATCTGGGACATC	8502
QY	3565	TGCGGAGCTGTGCTCTTACAGTACCAACCGCTTGAAGACTTAAGCTTGAATTAACGA	3624

Db	8503	TCGGACCGCTGTGCCCTTTAGCTACCA	CCGCTTGAGACATTACTCTTGATTGA	ACGA	8562
Qy	3625	GGATTGTGAACTTCTGGACGCGAGGGG	GTGGGAGCCCTCAAAATTGGTGAATC	CC	3684
Db	8563	GGATTGTGAACTTCTGGACGCGAGGGG	GTGGGAGCCCTCAAAATTGGTGAATC	CC	8622
Qy	3685	TACAGTATTGGAGTCACGAACTAAGA	ATAAGTAGTCTGTAACTTGTCAATG	CCACAGCCA	3744
Db	8623	TACAGTATTGGAGTCACGAACTAAGA	ATAAGTAGTCTGTAACTTGTCAATG	CCACAGCCA	8682
Qy	3745	TACAGTAGCTGAGGGGACAGATAGGG	TTATAGAAATATTACAAAGAGCTTT	TACAGCTA	3804
Db	8683	TAGCAGTAGCTGAGGGGACAGATAGGG	TTATAGAAATATTACAAAGAGCTTT	TACAGCTA	8742
Qy	3805	TTGCCCATACCTACCTAGAAAGATTA	AGACAGGCTTGGAAAGAAATTTTGTAT	AAATGGGT	3864
Db	8743	TTGCCCATACCTACCTAGAAAGATTA	AGACAGGCTTGGAAAGAAATTTTGTAT	AAATGGGT	8802
Qy	3865	GGCAAGTGGTCAAAAAGTACTGTGAT	TGGATTCGCTGTAAAGGAAAGAAAT	GAGACGA	3924
Db	8803	GGCAAGTGGTCAAAAAGTACTGTGAT	TGGATTCGCTGTAAAGGAAAGAAAT	GAGACGA	8862
Qy	3925	GCTAGCCACGACGAGATGGGGTGGG	AGCGATTTCTCGAGATCTAGA	3971	
Db	8863	GCTAGCCACGACGAGATGGGGTGGG	AGCGATTTCTCGAGATCTAGA	8909	
RESULT 7					
US-10-000-511A-2					
; Sequence 2, Application US/10000511A					
; Publication No. US20030158131A1					
; GENERAL INFORMATION:					
; APPLICANT: Aldovini, Anna					
; TITLE OF INVENTION: DNA VECTORS CONTAINING MUTATED HIV PROVIRUSES					
; FILE REFERENCE: 101353-165					
; CURRENT APPLICATION NUMBER: US/10/000,511A					
; CURRENT FILING DATE: 2001-10-23					
; PRIOR APPLICATION NUMBER: 60/242,589					
; PRIOR FILING DATE: 2000-10-23					
; PRIOR APPLICATION NUMBER: 60/243,432					
; PRIOR FILING DATE: 2000-11-28					
; NUMBER OF SEQ ID NOS: 37					
; SOFTWARE: PatentIn Ver. 2.1					
; SEQ ID NO 2					
; LENGTH: 9719					
; TYPE: DNA					
; ORGANISM: Human Immunodeficiency virus					
; FEATURE:					
; NAME/KEY: CDS					
; LOCATION: (1)..(9717)					
US-10-000-511A-2					
Query Match					
Best Local Similarity 41.6%; Score 2593; DB 12; Length 9719;					
Matches 2636; Conservative 0; Mismatches 45; Indels 6; Gaps 1;					
Qy	1291	CCATGAGAGTGAAGGAGAGTAATCAGAC	ACTTGTGGAGATGGGGGTGGAATGGGCAC	CA	1350
Db	6223	CAATGAGAGTGAAGGAGAAATATCAGAC	ACTTGTGGAGATGGGGGTGGAATGGGCAC	CA	6282
Qy	1351	TGCTCTTGGGATATTGATGATCTGATAG	TGCTTACAGAAAAATTGTGGGCTACACG	CTATT	1410
Db	6283	TGCTCTTGGGATTTGATGATCTGATAGT	CTTACAGAAAAATTGTGGGCTACAGCTAT	TT	6342
Qy	1411	ATGGGGTACCTGTGTGGAAGAACCAAC	CACTTAATTTTGTGATAGATGCTAAAG		1470
Db	6343	ATGGGGTACCTGTGTGGAAGAACCAAC	CACTTAATTTTGTGATAGATGCTAAAG		6402
Qy	1471	CATATGATACAGAGGTGCATATATGTTT	GGGCCACACATGCGGTGTACACAGACCC	CA	1530
Db	6403	CATATGATACAGAGGTGCATATATGTTT	GGGCCACACATGCGGTGTACACAGACCC	CA	6462
Qy	1531	ACCACAGAAGTAGTATTGTTGAATGTGA	CAAGAAAAATTTTAACATGTGGAAAAAT	GTACA	1590

|||||
Db 6463 ACCCACAGAGATGTAATGTGTAATGTGACAGAAATTTTAACATGTGGAAAAATACACA 6522
Qy 1591 TGGTAGAACAGATGTCATGAGGATATTAATCAGTTTATGGGATCAAAAGCCTAAAGCCATGTG 1650
Db 6523 TGGTAGAACAGATGTCATGAGGATATTAATCAGTTTATGGGATCAAAAGCCTAAAGCCATGTG 6582
Qy 1651 TAAAATTAACCCCATCTGTGTATAGTTTAAAGTCACCTGATTTGACAAATGATATCTATA 1710
Db 6583 TAAAATTAACCCCATCTGTGTATAGTTTAAAGTCACCTGATTTGACAAATGATATCTATA 6642
Qy 1711 CCAATAGTAGAGGGAGATGATTAATGGAGAAAGAGATAAAACGTCCTTCA 1770
Db 6643 CCAATAGTAGAGGGAGATGATTAATGGAGAAAGAGATAAAACGTCCTTCA 6702
Qy 1771 ATATCAGACAGAGCATTAAGATTAAGTGCAGAAAGATATGCAATTCCTTTTAAACTTG 1830
Db 6703 ATATCAGACAGAGCATTAAGATTAAGTGCAGAAAGATATGCAATTCCTTTTAAACTTG 6762
Qy 1831 ATATAGTACCAATATGATTA-----TACCAAGCTATAGGTTGATATAGTTTGAACCTCAG 1884
Db 6763 ATATTAATTAACCAATATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 6822
Qy 1885 TCATTAACAGAGCCTGTCCAAAGGATCCTTTGAGCCAAATTCCTCATATATATGTGCC 1944
Db 6823 TCATTAACAGAGCCTGTCCAAAGGATCCTTTGAGCCAAATTCCTCATATATATGTGCC 6882
Qy 1945 CGGCTGTTTGGCATCTTAAATATGATTAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 2004
Db 6883 CGGCTGTTTGGCATCTTAAATATGATTAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 6942
Qy 2005 CAAATGTCAGACAGTACATATGACATGGAATCAGGCCAGTATCAATCACTCACTGC 2064
Db 6943 CAAATGTCAGACAGTACATATGACATGGAATCAGGCCAGTATCAATCACTCACTGC 7002
Qy 2065 TGTTAATGGAGCTGTGAGAGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2124
Db 7003 TGTTAATGGAGCTGTGAGAGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7062
Qy 2125 ATGCTAAACCATATATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2184
Db 7063 ATGCTAAACCATATATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7122
Qy 2185 ACAACATATACAGAAAAAGTATCCGTATCCAGAGGGGACCAAGGAGACATTTGTTCAA 2244
Db 7123 ACAACATATACAGAAAAAGTATCCGTATCCAGAGGGGACCAAGGAGACATTTGTTCAA 7182
Qy 2245 TAGGAAAAATAGGAAAAATGAGAACACATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2304
Db 7183 TAGGAAAAATAGGAAAAATGAGAACACATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 7242
Qy 2305 CCACTTTAAACAGATAGTAGCAAAATTAAGAGAACATTTGAAATTAATTAATTAATTAATTAATTA 2364
Db 7243 ACACTTTAAACAGATAGTAGCAAAATTAAGAGAACATTTGAAATTAATTAATTAATTAATTAATTA 7302
Qy 2365 TCTTTAAGCAATCTCGAGAGGGGACCCAGAAATTTGAACGACAGTTTAAATTTGAGAG 2424
Db 7303 TCTTTAAGCAATCTCGAGAGGGGACCCAGAAATTTGAACGACAGTTTAAATTTGAGAG 7362
Qy 2425 GGGAAATTTCTACTGTAATTAACACACAGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2484
Db 7363 GGGAAATTTCTACTGTAATTAACACACAGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 7422
Qy 2485 GGAATAGTAGAGGCTCAAAATTAACACTGAAGAGTGCACAAATCAGCTCCCATGAGAA 2544
Db 7423 GGAATAGTAGAGGCTCAAAATTAACACTGAAGAGTGCACAAATCAGCTCCCATGAGAA 7482
Qy 2545 TAAAACATTTATTAACATGTGGCAGAGATAGGAAAAAGCAATGTATGCCCTCCATCA 2604
Db 7483 TAAAACATTTATTAACATGTGGCAGAGATAGGAAAAAGCAATGTATGCCCTCCATCA 7542
Qy 2605 GTGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2664
|||||

Db 7543 GTGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7602
Qy 2665 ATAAACAATAGGCTCCAGAGATCTTCAGACCTGGAGAGGCGATATAGGAGCAATTTGA 2724
Db 7603 ATAGCAACAATAGGCTCCAGAGATCTTCAGACCTGGAGAGGAGATATAGGAGCAATTTGA 7662
Qy 2725 GAAATGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2784
Db 7663 GAAATGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7722
Qy 2785 AGGCAAGAGAGAGTGTGTCAGAGAGAAAAAGAGAGAGGGAATTTGAGAGCTTTGTTCC 2844
Db 7723 AGGCAAGAGAGAGTGTGTCAGAGAGAAAAAGAGAGAGGGAATTTGAGAGCTTTGTTCC 7782
Qy 2845 TTGGGTTCTTGGAGCAGAGAGGAGCACTATGGGCTGCACAGTCAATGACCTGACGGTAC 2904
Db 7783 TTGGGTTCTTGGAGCAGAGAGGAGCACTATGGGCTGCACAGTCAATGACCTGACGGTAC 7842
Qy 2905 AGGCCAGACATTTATTTGTCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2964
Db 7843 AGGCCAGACATTTATTTGTCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7902
Qy 2965 AGGCGAACAGCATCTGTGCAACTGCAGCTGAGGGGCAATCAACAGCTCAGGCAAGAA 3024
Db 7903 AGGCGAACAGCATCTGTGCAACTGCAGCTGAGGGGCAATCAACAGCTCAGGCAAGAA 7962
Qy 3025 TCCCTGGCTGTGAAAGATTAACCTTAAGGATCAACAGCTCCTGGGGATTTGGGGTTCCTG 3084
Db 7963 TCCCTGGCTGTGAAAGATTAACCTTAAGGATCAACAGCTCCTGGGGATTTGGGGTTCCTG 8022
Qy 3085 GAAATCTCATTTGACACCTGCTGTGCTTTGGAATGCTAGTTGAGTAAATTAATTTCTG 3144
Db 8023 GAAATCTCATTTGACACCTGCTGTGCTTTGGAATGCTAGTTGAGTAAATTAATTTCTG 8082
Qy 3145 AACAGATTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3204
Db 8083 AACAGATTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8142
Qy 3205 GCTTAATTAACATCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3264
Db 8143 GCTTAATTAACATCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8202
Qy 3265 TATTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3324
Db 8203 TATTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8322
Qy 3325 GGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3384
Db 8263 GGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8322
Qy 3385 CTGTACTTTCTATAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3444
Db 8323 CTGTACTTTCTATAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8382
Qy 3445 ACCTCCCAATCCCGAGGGGACCCGACAGGCCCCGAGAAATTAAGAAAGAGGTGGAGAGA 3504
Db 8383 ACCTCCCAATCCCGAGGGGACCCGACAGGCCCCGAGAAATTAAGAAAGAGGTGGAGAGA 8442
Qy 3505 GAGACAGAGACAGATCCATTCGATTTGATTAAGAGGATCTTAAGACATTAATTTGGAGAGATC 3564
Db 8443 GAGACAGAGACAGATCCATTCGATTTGATTAAGAGGATCTTAAGACATTAATTTGGAGAGATC 8502
Qy 3565 TGGGAGAGCCTGTGCTCTTCCAGCTACACCGCTTGAGAGAGATTAATTAATTAATTAATTAATTAATTA 3624
Db 8503 TGGGAGAGCCTGTGCTCTTCCAGCTACACCGCTTGAGAGAGATTAATTAATTAATTAATTAATTAATTA 8562
Qy 3625 GGAATTTGGAGACTTTGGGAGCGAGGGGTTGGAGAGCCCTCAAAATTTGGTGGAAATCTCC 3684
Db 8563 GGAATTTGGAGACTTTGGGAGCGAGGGGTTGGAGAGCCCTCAAAATTTGGTGGAAATCTCC 8622
Qy 3685 TACAGATTTGAGATCAGAGCACTAAAGAAATGCTGTATTAATTTGCTCAATGGCCAGAGCA 3744
Db 8623 TACAGATTTGAGATCAGAGCACTAAAGAAATGCTGTATTAATTTGCTCAATGGCCAGAGCA 8682
|||||

```
OY 3745 TAGCAGTAGCTGAGGGGACACATAGGCTTATAGAAGTATTACAGACCTTATAGAGCTA 3804
      |||||||
Db 8663 TAGAGTAGCTGAGGGGACACATAGGCTTATAGAAGTATTACAGACCTTATAGAGCTA 8742
OY 3805 TTCCGCACATACCTAGAAAGTAAGACAGGGCTTGGAAGAATTTTGGTATAGATGGGT 3864
      |||||||
Db 8743 TTCCGCACATACCTAGAAAGTAAGACAGGGCTTGGAAGAATTTTGGTATAGATGGGT 8802
OY 3865 GGCAAGTGTCAAAAAAGTAGTGTGATTGGATGGCTGCTPAAGGGAAGAATGAGACGA 3924
      |||||||
Db 8803 GGCAAGTGTCAAAAAAGTAGTGTGATTGGATGGCTGCTPAAGGGAAGAATGAGACGA 8862
OY 3925 GCTAGCCACGACGAGATGGGGTGGAGACATATCTGAGATCTAGA 3971
      |||||||
Db 8863 GCTAGCCACGACGAGATGGGGTGGAGACATCTCGAGACCTGGA 8909

RESULT 8
US-10-286-332A-14
; Sequence 14, Application US/10286332A
; Publication No. US20030138459A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: METHOD OF VACCINATION THROUGH SPROTYPE ROTATION
; FILE REFERENCE: 22488-738
; CURRENT APPLICATION NUMBER: US/10/286,332A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 10/003,035
; PRIOR FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 3157
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Env/Tat/Rev
US-10-286-332A-14

Query Match 41.6%; Score 2592.4; DB 12; Length 3157;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 41; Indels 6; Gaps 1;
```

```
Db 838 TAAATTTAACCCCACTCTGTTTACTTTAAAGTGCACCTGATTTTGAAGATATATCTATA 897
      |||||||
OY 1711 CCAATAGTAGTAGGGGGGAATGATTAATGAGAAAGAGATAAAAAGCTCTTTCA 1770
      |||||||
Db 898 CCAATAGTAGTAGGGGGGAATGATTAATGAGAAAGAGATAAAAAGCTCTTTCA 957
OY 1771 ATATCAGCACAAGCATTAAGAGATTAAGTGCAGAAAAGTAATGCAATCTTTTATAACTTGG 1830
      |||||||
Db 958 ATATCAGCACAAGCATTAAGAGATTAAGTGCAGAAAAGTAATGCAATCTTTTATAACTTGG 1017
OY 1831 ATATAGTACCAATGATATA-----TACCACTTATAGGTATGTAACCTTCAG 1884
      |||||||
Db 1018 ATATATATCCAAATGATATATATATCTACCACTTATAGGTATGTAACCTTCAG 1077
OY 1885 TCATTTACACAGGCTGTCCAAAGGTATCTTTGAGCCAAATTCATACATTTATTTGGCCC 1944
      |||||||
Db 1078 TCATTTACACAGGCTGTCCAAAGGTATCTTTGAGCCAAATTCATACATTTATTTGGCCC 1137
OY 1945 CGGCTGTTTGGCATTTCTTAATATGTATTAATAGAGCTTCAATGGAACAGACCATGTA 2004
      |||||||
Db 1138 CGGCTGTTTGGCATTTCTTAATATGTATTAATAGAGCTTCAATGGAACAGACCATGTA 1197
OY 2005 CAAATGTCAGCAGATACATGTACACATGGGAATCAGGCGAGTAGTATCACTCACTGC 2064
      |||||||
Db 1198 CAAATGTCAGCAGATACATGTACACATGGGAATTAAGGCCAGTAGTATCACTCACTGC 1257
OY 2065 TGTTAATGGCAGCTACAGAGAAGATGTAGTAATTAATGATGTGCCAATTTACAGACA 2124
      |||||||
Db 1258 TGTTAATGGCAGCTACAGAGAAGATGTAGTAATTAATGATGTGCCAATTTACAGACA 1317
OY 2125 ATGCTAAACCATATATCTAGAGCTGAACACATCTGTGAATTAATGTTCAAGACCCA 2184
      |||||||
Db 1318 ATGCTAAACCATATATCTAGAGCTGAACACATCTGTGAATTAATGTTCAAGACCCA 1377
OY 2185 ACAACATACAGAAAAAGTATCCGTATCCAGAGGGGACAGGAGACATTTGTTACAA 2244
      |||||||
Db 1378 ACAACATACAGAAAAAGTATCCGTATCCGTATCCAGAGGGGACAGGAGACATTTGTTACAA 1437
OY 2245 TAGGAAAAATAGAAATATGAGACAGACATTTGATACATTAGTAGACCAAAATGGAATG 2304
      |||||||
Db 1438 TAGGAAAAATAGAAATATGAGACAGACATTTGATACATTAGTAGAGCAAAATGGAATA 1497
OY 2305 CCACCTTAAACAGATAGCTAGCAAAATTAAGAAACAATTTGAAATATATAACCAATRA 2364
      |||||||
Db 1498 ACACCTTAAACAGATAGATAGCAAAATTAAGAAACAATTTGAAATATATAACCAATRA 1557
OY 2365 TCTTTAAGCAATCCCTCAGAGAGGGGCCAGAAATTTGTAACGACAGTTTAAATTTGGAG 2424
      |||||||
Db 1558 TCTTTAAGCAATCCCTCAGAGAGGGGCCAGAAATTTGTAACGACAGTTTAAATTTGGAG 1617
OY 2425 GGGAAATTTTCTACTGTATTTCAACACACATGTTTAATAGTACTGTTTAAATAGTACTT 2484
      |||||||
Db 1618 GGGAAATTTTCTACTGTATTTCAACACACATGTTTAAATAGTACTGTTTAAATAGTACTT 1677
OY 2485 GGAGTACGAAGGGCAATTAACACTGAAGGAAGGACACATCACTCCATGCAAGAA 2544
      |||||||
Db 1678 GGAGTACGAAGGGCAATTAACACTGAAGGAAGGACACATCACTCCATGCAAGAA 1737
OY 2545 TAAACCAATTTATAAACATGTGCGAGAGTAGAGAAAGCAATGATCCCTCCATCA 2604
      |||||||
Db 1738 TAAACCAATTTATAAACATGTGCGAGAGTAGAGAAAGCAATGATCCCTCCATCA 1797
OY 2605 GTGCAAAATAGATGTTCAATCAATATTTACTGGGCTGCTATTAACAAGAGATGGTGTA 2664
      |||||||
Db 1798 GTGCAAAATAGATGTTCAATCAATATTTACTAGGGCTGCTATTAACAAGAGATGGTGTA 1857
OY 2665 ATTAACCAATATGGGTCCAGATCTTCAGACCTGGAAGGAGGAGATAGGAGCAATTTGGA 2724
      |||||||
Db 1858 ATAGCAAAATAGATGTTCAATCAATATTTACTAGGGCTGCTATTAACAAGAGATGGTGTA 1917
OY 2725 GAAAGTATTTATTAATATTAAGTAGAAAAATTTGAACCAATTTAGAGATACACCCACCA 2784
      |||||||
```

Db	1918	GAAGCAATTTATTAATAATATAAGTAAAAAATTTGAACCATTTAGAGGTAGCACCCACACA	1977
OY	2785	AGGCAAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGACAGTGGGAATATAGAGCTTTGTTCC	2844
Db	1978	AGGCAAAAGAGAAGTGGTGCAGAGAGAAAAAAGACAGTGGGAATATAGAGCTTTGTTCC	2037
OY	2845	TTGGGTTCTTGGAGACAGACAGGAAGACATATAGGGCTGGACGTCATATGACGCTGACGGTAC	2904
Db	2038	TTGGGTTCTTGGAGACAGACAGGAAGACATATGGGCGCAGGCTCATATGACGCTGACGGTAC	2097
OY	2905	AGGCGACAGCAATTAATTTGTTGATATATAGTCAGCAGACAGAACAAATTTGCTGAGGGCTATTG	2964
Db	2098	AGGCGACAGCAATTAATTTGTTGCTGTATAGTCAGCAGCAGAACAAATTTGCTGAGGGCTATTG	2157
OY	2965	AGGGGCAACACACATCTGTGTGAACCTCACAGTCGTGGGCACTAAACAGCTCCAGGCCAAGAA	3024
Db	2158	AGGGGCAACACACATCTGTGTGAACCTCACAGTCGTGGGCACTAAACAGCTCCAGGCCAAGAA	2217
OY	3025	TCCTGGGCTGTGGAAAGATACCTTAAGAGATCAACAGCTCTCTGGGGATTTGGGGTTGCTCTG	3084
Db	2218	TCCTGGGCTGTGGAAAGATACCTTAAGAGATCAACAGCTCTCTGGGGATTTGGGGTTGCTCTG	2277
OY	3085	GAAACATCATTTTGACACCACTGCTGTGCTTGGAATGCTAGTTGGAGATATTAATCTCTGG	3144
Db	2278	GAAACATCATTTTGACACCACTGCTGTGCTTGGAATGCTAGTTGGAGATATTAATCTCTGG	2337
OY	3145	AACAGATTTGGAAATACATGACCTGAGTGGAGTGGGACAGAGAAATTAACAAATTAACAA	3204
Db	2338	AACAGATTTGGAAATACATGACCTGAGTGGAGTGGGACAGAGAAATTAACAAATTAACAA	2397
OY	3205	GCTTAATTAACATCTCTTAATTAAGAAATCGCAAAACACAGCAAGAAAGATGACACAGAT	3264
Db	2398	GCTTAATTAACATCTCTTAATTAAGAAATCGCAAAACACAGCAAGAAAGATGACACAGAT	2457
OY	3265	TATTGAAATTTGATTAATGAGGACAGTTTGTGGAATTTGGTTAATACATTAACAAATTTGGGT	3324
Db	2458	TATTGAAATTTGATTAATGAGGACAGTTTGTGGAATTTGGTTAATACATTAACAAATTTGGGT	2517
OY	3325	GGTATATTAATAATTAATCATTAATGATAGGAGGCTTGGTAGGTTTAAGAAATAGTTTGTG	3384
Db	2518	GGTATATTAATAATTTATCATATGATAGGAGGCTTGGTAGGTTTAAGAAATAGTTTGTG	2577
OY	3385	CTGACATTTTCATATGTGATGATGATTTTGGGACGGGATTAATCAACATTTTTCGTTACAGACC	3444
Db	2578	CTGACATTTTCATATGTGATGATGATTTTGGGACGGGATTAATCAACATTTTTCGTTACAGACC	2637
OY	3445	ACCTCCCAATCCGAGGGGACCCGACAGGCCCGCAAGGAATTAAGAAAGAGGTTGGAAGAA	3504
Db	2638	ACCTCCCAATCCGAGGGGACCCGACAGGCCCGCAAGGAATTAAGAAAGAGGTTGGAAGAA	2697
OY	3505	GAGACACAGACAGATCCATTCGATTCGATTAAGAACGATCCTTAGACATTAATCTGGGACATC	3564
Db	2698	GAGACACAGACAGATTCATTCGATTAAGAACGATCCTTAGACATTAATCTGGGACATC	2757
OY	3565	TGCGAGAGCTGTGCTCTTCAGCTACCAACGCTTGAAGAGACTTAATCTCTGTTGTTAAGCA	3624
Db	2758	TGCGAGAGCTGTGCTCTTCAGCTACCAACGCTTGAAGAGACTTAATCTCTGTTGTTAAGCA	2817
OY	3625	GGATTTGTGGAACCTTCTGGAGACAGGCGGGGTGGAGGCCCTCAAAATTAATTTGTTGATCTCC	3684
Db	2818	GGATTTGTGGAACCTTCTGGAGACAGGCGGGGTGGAGGCCCTCAAAATTAATTTGTTGATCTCC	2877
OY	3685	TACAGTATTGAGACAGGAAGCTAAAGAAATGTGCTGTAACTGCTCAATGCGCACACAGCA	3744
Db	2878	TACAGTATTGAGACAGGAAGCTTAAGAAATGTGCTGTAACTGCTCAATGCGCACACAGCA	2937
OY	3745	TACAGTATAGCTGAGGGGACAGATATGGGTTATAGAAATTAATTAACAAGAGCTTATAGACTTA	3804
Db	2938	TACAGTATAGCTGAGGGGACAGATATGGGTTATAGAAATTAATTAACAAGAGCTTATAGACTTA	2997
OY	3805	TTTGCCCACTATACCTAGAAAGAAATTAAGACAGGGCTTTGGAAAGGATTTTCTATATAGATGGGT	3864
Db	2998	TTTGCCCACTATACCTAGAAAGAAATTAAGACAGGGCTTTGGAAAGGATTTTCTATATAGATGGGT	3057

```

RESULT 9
US-10-003-035-14
; Sequence 14, Application US/10003035
; Publication NO. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: SeqId version 3.1
; SEQ ID NO 14
; LENGTH: 3157
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Env/Tac/Rev
US-10-003-035-14

```

Query Match	Best Local Similarity	41.6% ; Score 2592.4 ;	DB 13 ; Length 3157 ;
Matches 2633 ;	Conservative 0 ; Mismatches 41 ; Indels 6 ; Gaps 1 ;		
QY	1291	CCATGAGAGTGAAGAGAAAGTATCAGCACCTTGAGAGATGGGGGTGGAATGGGGCCACCA	1350
Db	478	CAATGAGAGTGAAGAGAAATATCAGCACCTTGAGAGATGGGGGTGAGATGGGGCCACCA	537
QY	1351	TGCTCCTTGGGATTTATGATCTGTGTAGTGTCTACAGAAAAATGTGGGTACCGTCTATT	1410
Db	538	TGCTCCTTGGGATTTATGATCTGTGTAGTGTCTACAGAAAAATGTGGGTACAGTCTATT	597
QY	1411	ATGGGGTACCTGTGTGAGGAAGCAACCAACCTCATTTTGTGCTACAGATGCTTAAG	1470
Db	598	ATGGGGTACCTGTGTGAGGAAGCAACCAACCTCATTTTGTGCTACAGATGCTTAAG	657
QY	1471	CATATGATACAGAGGTACATATGTTTTGGGGCCACACATGCGCTGTATCCCAAGACCCCA	1530
Db	658	CATATGATACAGAGGTACATATGTTTTGGGGCCACACATGCGCTGTATCCCAAGACCCCA	717
QY	1531	ACCCACAAGATGATTTGGTAAATGTGACAGAAAAATTTTAACATGTGAAAAATGACA	1590
Db	718	ACCCACAAGATGATTTGGTAAATGTGACAGAAAAATTTTAACATGTGAAAAATGACA	777
QY	1591	TGTGTGAACAAGATGCATAGAGATATATTCAGTTTATGGGATTCAAAAGCTTAACCCATGTG	1650
Db	778	TGTGTGAACAAGATGCATAGAGATATATTCAGTTTATGGGATTCAAAAGCTTAACCCATGTG	837
QY	1651	TAAAAATTAACCCCACTGTGTAGTTTAAAGTGCATGTATTTGAAGAATGATACTAATA	1710
Db	838	TAAAAATTAACCCCACTGTGTAGTTTAAAGTGCATGTATTTGAAGAATGATACTAATA	897
QY	1711	CCAACTAGTAGAGGGGAGAAATGATAATGAGAAAGAGAGATAAAAAACTGCTCTTTCA	1770
Db	898	CCAACTAGTAGAGGGGAGAAATGATAATGAGAAAGAGAGATAAAAAACTGCTCTTTCA	957
QY	1771	ATATCAGCACAGCATTAAGATTAAGTGCAGAAAGATATGCAATTCCTTTTATTAACCTTG	1830
Db	958	ATATCAGCACAGCATTAAGATTAAGTGCAGAAAGATATGCAATTCCTTTTATTAACCTTG	1017


```
RESULT 10
US-10-286-332A-22
; Sequence 22, Application US/10286332A
; Publication No. US20030138459A1
; GENERAL INFORMATION:
; APPLICANT: Manu, Danher
; TITLE OF INVENTION: METHOD OF VACCINATION THROUGH SEROTYPE ROTATION
; FILE REFERENCE: 22488-738
; CURRENT APPLICATION NUMBER: US/10/286,332A
; PRIOR FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 10/003,035
; PRIOR FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 2950
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env/Nef from strain BH10
US-10-286-332A-22

Query Match      41.4%; Score 2576; DB 12; Length 2950;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 2635; Conservative 0; Mismatches 45; Indels 12; Gaps 2;

QY      1286  CGCCACCATGAGATGGAAGGAAGATCAGCAGCTTGTGAGATGGGGGTGGAATGGGG 1345
DB      6      CGCCACCATGGAGTGAAGGAATATCAGCAGCTTGTGAGATGGGGGTGGAATGGGG 65
QY      1346  CACCATGCTCCTGGGATGATGATCTGTGATGCTACAGAAAAATTTGGTGCACCGT 1405
DB      6      CACCATGCTCCTGGGATGATGATCTGTGATGCTACAGAAAAATTTGGTGCACCGT 125
QY      1406  CTATTATGGGGTACCTGTGTGGAAGGAAGCAACCACTCTATTATTGTGCATCAGATGC 1465
DB      126     CTATTATGGGGTACCTGTGTGGAAGGAAGCAACCACTCTATTATTGTGCATCAGATGC 185
QY      1466  TAAAGCATATGATACAGAGGTACATATGTTGGGCCACACATCCCTGTGTACCCACAGA 1525
DB      186     TAAAGCATATGATACAGAGGTACATATGTTGGGCCACACATCCCTGTGTACCCACAGA 245
QY      1526  CCCCACCCCAAGAAGATGATGTAATGTGACAGAAAAATTTAAACATGTGGAATAA 1585
DB      246     CCCCACCCCAAGAAGATGATGTAATGTGACAGAAAAATTTAAACATGTGGAATAA 305
QY      1586  TGACATGATAGAACAGATGCATGAGATATATCATGTTATGGGATCAAAAGCCTTAAAGCC 1645
DB      306     TGACATGATAGAACAGATGCATGAGATATATCATGTTATGGGATCAAAAGCCTTAAAGCC 365
QY      1646  ATGCTGTAATTAACCCCACTCTGTGTAGTTTAAAGTGCACCTGATTGGAAGATGATAC 1705
DB      366     ATGCTGTAATTAACCCCACTCTGTGTAGTTTAAAGTGCACCTGATTGGAAGATGATAC 425
QY      1706  TAATACCAATATGATGAGGAGGAGATGATATGATGAGAGAAAGAGATTAATAAATCTGTC 1765
DB      426     TAATACCAATATGATGAGGAGGAGATGATATGATGAGAGAAAGAGATTAATAAATCTGTC 485
QY      1766  TTTCATATATGAGCACAAGCATATAGATTAAGGTGCAGAAAAATATGATCTTTTATTA 1825
DB      486     TTTCATATATGAGCACAAGCATATAGATTAAGGTGCAGAAAAATATGATCTTTTATTA 545
QY      1826  ACTTGATATAGTACCATATAGTAA-----TACCAAGCTATATGTTGATAGTTTAAAC 1879
DB      546     ACTTGATATATATACCAATATAGTAACTACACAGCTATACGTTGACAAAGTTGTAACAC 605
QY      1880  CTCAGTATTAACAGAGCGCTGTCCAAAGGTATCCTTTGAGCCATTTCCCATACATTTATG 1939
DB      606     CTCAGTATTAACAGAGCGCTGTCCAAAGGTATCCTTTGAGCCATTTCCCATACATTTATG 665
```

```
QY      1940  TGCCCGCGCTGTTTGCAGATTTCTAAATGTAATATAGACGTTCAATGGAACAGGACC 1999
DB      666     TGCCCGCGCTGTTTGCAGATTTCTAAATGTAATATAGACGTTCAATGGAACAGGACC 725
QY      2000  ATGTACAATATGTCAGCAGATGACATATGATGATGATGATGATGATGATGATGATGATG 2059
DB      726     ATGTACAATATGTCAGCAGATGACATATGATGATGATGATGATGATGATGATGATGATG 785
QY      2060  ACTGCTTTAAATGAGCTGTAGCAGAAGAGTGTATATGATGATGATGATGATGATGATGAT 2119
DB      786     ACTGCTTTAAATGAGCTGTAGCAGAAGAGTGTATATGATGATGATGATGATGATGATGATG 845
QY      2120  AGACAATGCTAAACCAATATATGATGATGATGATGATGATGATGATGATGATGATGATG 2179
DB      846     AGACAATGCTAAACCAATATATGATGATGATGATGATGATGATGATGATGATGATGATG 905
QY      2180  ACCCAACACATATCAAGAAAAAGTATCCGTATCCAGAGGGGACAGGAGAGCATTTGT 2239
DB      906     ACCCAACACATATCAAGAAAAAGTATCCGTATCCAGAGGAGACAGGAGAGCATTTGT 965
QY      2240  TACATATGAAAAATAGAAAAATGAGACAAGCATTTGATGATGATGATGATGATGATGATG 2299
DB      966     TACATATGAAAAATAGAAAAATGAGACAAGCATTTGATGATGATGATGATGATGATGATG 1025
QY      2300  GAATGCCACTTTAAACAGATAGCTACCAATTTAAGAGAACAAATTTGGAATATATAAAC 2359
DB      1026     GAATGAACACTTTAAACAGATAGCTACCAATTTAAGAGAACAAATTTGGAATATATAAAC 1085
QY      2360  AATATCTTTAAGCAATCTCAGAGGGGACCCAGAAATTTGTAACGACAGTTTAAATG 2419
DB      1086     AATATCTTTAAGCAATCTCAGAGGGGACCCAGAAATTTGTAACGACAGTTTAAATG 1145
QY      2420  TGAAGGGGAATTTTCTACGTATATCAACAACCTTTAATATGATGATGATGATGATGATG 2479
DB      1146     TGAAGGGGAATTTTCTACGTATATCAACAACCTTTAATATGATGATGATGATGATGATG 1205
QY      2480  TACTTGAATAGTGAAGGGTCAATTAACACTGAAAGAGTGAACAATCACTCCATG 2539
DB      1206     TACTTGAATAGTGAAGGGTCAATTAACACTGAAAGAGTGAACAATCACTCCATG 1265
QY      2540  CAGATTAACCAATTTTAAACATGTCGAGAGAGTGAAGAGAGTGAAGAGTGAAGAGTGA 2599
DB      1266     CAGATTAACCAATTTTAAACATGTCGAGAGAGTGAAGAGAGTGAAGAGTGAAGAGTGA 1325
QY      2600  CATCAGTGAACAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2659
DB      1326     CATCAGTGAACAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1385
QY      2660  TGCTAATTAACCAATGGGTCGAGATCTTCAGACCTGAGAGGAGCGATATGAGGACAA 2719
DB      1386     TGCTAATTAACCAATGATGTCGAGATCTTCAGACCTGAGAGGAGCGATATGAGGACAA 1445
QY      2720  TTGAGAGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2779
DB      1446     TTGAGAGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1505
QY      2780  CACCAAGGCAAGAGAGAGTGTGAGAGAGAAAAAGAGAGAGTGGGAATGAGAGCTT 2839
DB      1506     CACCAAGGCAAGAGAGAGTGTGAGAGAGAAAAAGAGAGAGTGGGAATGAGAGCTT 1559
QY      2840  GTTCCTTGGGTTCTTGGGAGCAGAGAGACATATGGGCTGCACGTCAATGACGCTGAC 2899
DB      1560     GTTCCTTGGGTTCTTGGGAGCAGAGAGACATATGGGCTGCACGTCAATGACGCTGAC 1619
QY      2900  GGTACAGGCCAGACATATATGCTGATATATGTCAGACAGACAGACATTTGCTGAGGGC 2959
DB      1620     GGTACAGGCCAGACATATATGCTGATATATGTCAGACAGACAGACATTTGCTGAGGGC 1679
QY      2960  TATGAGGGCCAGACAGATGCTGTGCAACGTCACAGTCTGGGGCATCAACAGCTCCAGGC 3019
DB      1680     TATGAGGGCCAGACAGATGCTGTGCAACGTCACAGTCTGGGGCATCAACAGCTCCAGGC 1739
```


QY	3020	AAGAATCCTGGCTGTGGAAAGATATCCATAAAGATCAACAGCTCCTGGGATTTGGGGTGG	3079
Db	1740	AAGATTCCTGGCTGTGGAAAGATATCCATAAAGATCAACAGCTCCTGGGATTTGGGGTGG	1799
QY	3080	CTCTGAGAAACCTCATTTTGCACACAGCTGCTGCTCTGGAAATCTAGTTGGAGTAATTAATC	3139
Db	1800	CTCTGAGAAACCTCATTTTGCACACAGCTGCTGCTCTGGAAATCTAGTTGGAGTAATTAATC	1859
QY	3140	TCCTGGAACAGATTTGGAAATAACATGACCTGGATGGAGTGGGACAGAAATTAACAATTA	3199
Db	1860	TCCTGGAACAGATTTGGAAATAACATGACCTGGATGGAGTGGGACAGAAATTAACAATTA	1919
QY	3200	CACAAAGCTTAATACACTTCCTTAATTTGAAGAATTCGCAAAACGACAGAAAAAGATGAACA	3259
Db	1920	CACAAAGCTTAATACACTTCCTTAATTTGAAGAATTCGCAAAACGACAGAAAAAGATGAACA	1979
QY	3260	AGAATTTATGGAATTTAGATTAATGGGACAGTTTGGAAATGGTTTAACATAACAATTTG	3319
Db	1980	AGAATTTATGGAATTTAGATTAATGGGACAGTTTGGAAATGGTTTAACATAACAATTTG	2039
QY	3320	GCTGTGGTATATAAAATTAATTCATATATGATAGTAGAGAGGCTTGGTAGGTTTAAAGAAAT	3379
Db	2040	GCTGTGGTATATAAAATTAATTCATATATGATAGTAGAGAGGCTTGGTAGGTTTAAAGAAAT	2099
QY	3380	TTTTTGCCTGACTTTCCTATAGTAATGAGATTAGCCAGGAGATTTCCACATTATCGTTTCA	3439
Db	2100	TTTTTGCCTGACTTTCCTATAGTAATGAGATTAGCCAGGAGATTTCCACATTATCGTTTCA	2159
QY	3440	GACCACCTCCCATCCCGAGGGGACCCGACAGGCCGCAAGAGATTAAGAAAGAAAGTGG	3499
Db	2160	GACCACCTCCCATCCCGAGGGGACCCGACAGGCCGCAAGAGATTAAGAAAGAAAGTGG	2219
QY	3500	AGAGAGAGACAGAGACAGATCCATTCGATTAATGATAACGAGATCCTTAGCACTTATCTGGGA	3559
Db	2220	AGAGAGAGACAGAGACAGATCCATTCGATTAATGATAACGAGATCCTTAGCACTTATCTGGGA	2279
QY	3560	CGATCTGCGGAGCCTGTGCTCTTCACGTACACCCGCTTGAGAGACTTACTCTTGATTTGT	3619
Db	2280	CGATCTGCGGAGCCTGTGCTCTTCACGTACACCCGCTTGAGAGACTTACTCTTGATTTGT	2339
QY	3620	AACGAGAGTTGTGAACTTCTGGGACGAGGGGTGGGAAGCCCTCAAAATTTGTGGAA	3679
Db	2340	AACGAGAGTTGTGAACTTCTGGGACGAGGGGTGGGAAGCCCTCAAAATTTGTGGAA	2399
QY	3680	TCCTCTCAGTATTTGGAGTCAAGAACTAAAGAAATAGTCTTTAACTTGTCTCAATGCCAC	3739
Db	2400	TCCTCTCAGTATTTGGAGTCAAGAACTAAAGAAATAGTCTTTAACTTGTCTCAATGCCAC	2459
QY	3740	AGCCATATGACGTAAGCTGAGGGGACAGATAGGGTTATAGAAAGTATTACAGCAGCTTATAG	3799
Db	2460	AGCCATATGACGTAAGCTGAGGGGACAGATAGGGTTATAGAAAGTATTACAGCAGCTTATAG	2519
QY	3800	AGCATTTGCGCACATTAACCTAGAAACAATAAAGACAGGCTTGGAAAGATTTTGCATTAAGA	3859
Db	2520	AGCATTTGCGCACATTAACCTAGAAACAATAAAGACAGGCTTGGAAAGATTTTGCATTAAGA	2579
QY	3860	TGGGTGGCAAGTGGTCAAAAAGTAGTGTGATTTGGATGGCCGTGTAAAGGAAAGAAATGA	3919
Db	2580	TGGGTGGCAAGTGGTCAAAAAGTAGTGTGATTTGGATGGCCGTGTAAAGGAAAGAAATGA	2639
QY	3920	GACGAGCTGAGCCAGCAGCAGATGGGGTGGGAGCAGATATCTCGAATCTTACA 3971	
Db	2640	GACGAGCTGAGCCAGCAGCAGATGGGGTGGGAGCAGATATCTCGAATCTTACA 2691	

	/	CURRENT APPLICATION NUMBER:	US/10/003_035	
	/	CURRENT FILING DATE:	2001-11-01	
	/	PRIOR APPLICATION NUMBER:	09/585,599	
	/	PRIOR FILING DATE:	2000-06-02	
	/	PRIOR APPLICATION NUMBER:	PCT/US01/18238	
	/	PRIOR FILING DATE:	2001-06-04	
	/	NUMBER OF SEQ ID NOS:	75	
	/	SOFTWARE:	PatentIn version 3.1	
	/	SEQ ID NO 22		
	/	LENGTH:	2950	
	/	TYPE:	DNA	
	/	ORGANISM:	Artificial sequence	
	/	FEATURE:		
	/	OTHER INFORMATION:	Modified Env/Nef from strain BH10	
	/	US-10-003-035-22		
Qy		Query Match	41.4%;	Score 2576; DB 13; Length 2950;
		Best Local Similarity	97.9%;	Pred. No. 0;
		Matches 2635;	Conservative	0; Mismatches 45; Indels 12; Gaps 2;
Qy		1286	CGCCACCACTGAGAGTGAAGAAGATATCACACTTGAGGATGGGGGTGAATGGGG	13445
		6	CGCCACCACTGGAGAGTGAAGAAGAAATATCAGCACCCTTGAGAGATGGGGGTGAGATGGGG	65
Qy		1346	CACCATGTCCTTGGGATATTGATGANTCTGATGCTACAGAAAAAATTTGGGGTCACCGT	14050
		66	CACCATGTCCTTGGGATGGTATGATGATCTGATGCTACAGAAAAAATTTGGGGTCACAGT	125
Qy		1406	CTATTATGGGGTACTCTGTGTGGAAAGAACCAACCACTCATTTTTGTGCATCAGATGC	14655
		126	CTATTATGGGGTACTCTGTGTGGAAAGAACCAACCACTCATTTTTGTGCATCAGATGC	185
Qy		1466	TAAACATATGATACAGAGATACATATATGTTGGGCCACACATGCGCTGTGACCCACAGA	15252
		186	TAAACATATGATACAGAGATACATATATGTTGGGCCACACATGCGCTGTGACCCACAGA	245
Qy		1526	CCCCAACCCACAGAAAGTAGTATTGGTAAATGTGACAGAAAAATTTTAACATGTGGAAAA	15858
		246	CCCCAACCCACAGAAAGTAGTATTGGTAAATGTGACAGAAAAATTTTAACATGTGGAAAA	305
Qy		1586	TGACATGTTAGAACAGATGCGATGAGAGATATATCAGTTTATGGGATCAAACCTTAAAGCC	16455
		306	TGACATGTTAGAACAGATGCGATGAGAGATATATCAGTTTATGGGATCAAACCTTAAAGCC	365
Qy		1646	ATGTGTAAAATTTAAACCCCACCTGTGTGTTAAAGTGCACCTGATTTGAAGATGATAC	17050
		366	ATGTGTAAAATTTAAACCCCACCTGTGTGTTAAAGTGCACCTGATTTGAAGATGATAC	425
Qy		1706	TAAATACAATAGTAGTAGCGGGAGAAATGATAATGAGAGAAAGAGATATAAAAACTGCTC	17655
		426	TAAATACAATAGTAGTAGCGGGAGAAATGATAATGAGAGAAAGAGATATAAAAACTGCTC	485
Qy		1766	TTTCAATATCAGCACAACCATTAAGAGATTAAGSTGCAGAAAGAAATATGCAATCTTTTATTA	18235
		486	TTTCAATATCAGCACAACCATTAAGAGATTAAGSTGCAGAAAGAAATATGCAATCTTTTATTA	545
Qy		1826	ACTTGATATAGTACCAATATAGATAA-----TACCAGCATATGTTGATTAAGTTTGAACAC	18797
		546	ACTTGATATATTAATACCAATATAGATAATGATACTACACAGCATATGCTTTGACAAAGTTGTAACAC	605
Qy		1880	CTCAGATCATTAACAGAGGCTGTCCAAGAGTATCCTTTGAGGCAATTTCCATATCATTTATG	19399
		606	CTCAGATCATTAACAGAGGCTGTCCAAGAGTATCCTTTGAGGCAATTTCCATATCATTTATG	665
Qy		1940	TGCCCCGGCTGGTTTTGCGATTCTTAAATGTAATATATTAAGCGTTCAATGGAACAGGACC	19999
		666	TGCCCCGGCTGGTTTTGCGATTCTTAAATGTAATATTAAGCGCTTCAATGGAACAGGACC	725
Qy		2000	ATGTACAATGTCAAGCAGATACATATGATGACATGAGATCGGGCAGATGATCAACGTCA	20599
		726	ATGTACAATGTCAAGCAGATACATATGATGACATGAGATTTAGGCCAGATGATCAACTCA	785
Qy		2060	ACTGCTGTAAATGGCACTTACGAGAAAGATGTAGTAATTAGATCTGCCAATTTTCAAC	21199

```
|||||
Db 786 ACTGCTGTAATAGGAGTGTGGCAGAAAGAGGTAGTAATGATCTGCCAATTCAC 845
Qy 2120 AGAATAAGCTAAACCCATAATAGTACAGCTGACACATCTGTATAAATTAATGTACAG 2179
Db 846 AGAATAAGCTAAACCCATAATAGTACAGCTGACACCAATCTGTATAAATTAATGTACAG 905
Qy 2180 ACCCAACACAAATACAGAAAAAGTATCCGTATCCAGAGGGACAGGAGAGCATTTGT 2239
Db 906 ACCCAACACAAATACAGAAAAAGTATCCGTATCCAGAGGAGACAGGAGAGCATTTGT 965
Qy 2240 TACATATAGAAAAATAGGAATATAGACAAACATTTGTATACATTAGTAGAGCAAAATG 2299
Db 966 TACATATAGAAAAATAGGAATATAGACAAACATTTGTATACATTAGTAGAGCAAAATG 1025
Qy 2300 GAATGCCACTTAAACAGATAGCTAGCAAAATTAAGAACAAATTTGAAATTAATAAAC 2359
Db 1026 GAATTAACACTTTAAACAGATAGATAGCAAAATTAAGAACAAATTTGAAATTAATAAAC 1085
Qy 2360 AATATCTTTAAGCAATCTCAGAGGGAGGCCAGAAATTTGTAAGCAGACAGTTTAAATG 2419
Db 1086 AATATCTTTAAGCAATCTCAGAGGGAGGCCAGAAATTTGTAAGCAGACAGTTTAAATG 1145
Qy 2420 TGGAGGGGAATTTTCTACTGTATTAATCAACAACTGTTTATATGTAAGTCTGTTTAAATG 2479
Db 1146 TGGAGGGGAATTTTCTACTGTATTAATCAACAACTGTTTATATGTAAGTCTGTTTAAATG 1205
Qy 2480 TACTTGGAGTACTGACGAGGTCATTAATACACTGAAGAGTAGACACATCACACTCCCATG 2539
Db 1206 TACTTGGAGTACTGACGAGGTCATTAATACACTGAAGAGTAGACACATCACACTCCCATG 1265
Qy 2540 CAGAAATAACAATTTATTAACATGTGGCAGAGAGTAGAGAAACCAATGTATGCCCTCC 2599
Db 1266 CAGAAATAACAATTTATTAACATGTGGCAGAGAGTAGAGAAACCAATGTATGCCCTCC 1325
Qy 2600 CATAGTGGACAAATTTAGATGTTCATCAAAATTTACTGGGCTGCTATTTAACAAAGATGG 2659
Db 1326 CATAGTGGACAAATTTAGATGTTCATCAAAATTTACTGGGCTGCTATTTAACAAAGATGG 1385
Qy 2660 TGGTAATTAACAAATAGGTCGAGATCTTCAGACCTGAGAGAGCGCATATGAGGGACAA 2719
Db 1386 TGGTAATTAACAAATAGGTCGAGATCTTCAGACCTGAGAGAGCGCATATGAGGGACAA 1445
Qy 2720 TTGGAGAGTGAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2779
Db 1446 TTGGAGAGTGAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1505
Qy 2780 CACCAAGGCAAAAGAGAGTGTGACAGAGAGAAAAAGACATGGGAATAGAGCTTT 2839
Db 1506 CACCAAGGCAAAAGAGAGTGTGACAGAGAGAAAAAGACATGGGAATAGAGAGCTTT 1559
Qy 2840 GTTCTTGGGTTCTTGGGAGCAGAGAGACACTATGGGCTGCACATTAATGAGCTGAC 2899
Db 1560 GTTCTTGGGTTCTTGGGAGCAGAGAGACACTATGGGCTGCACATTAATGAGCTGAC 1619
Qy 2900 GGTACAGGCGAGACAAATTTATTTGCTGATTAATGTCACAGAGAGAAATTTGCTGAGGGC 2959
Db 1620 GGTACAGGCGAGACAAATTTATTTGCTGATTAATGTCACAGAGAGAAATTTGCTGAGGGC 1679
Qy 2960 TATTGAGGCGACAGACATCTGTTGCACTCAGCTGTGGGGCATCAACAGCTCCAGGC 3019
Db 1680 TATTGAGGCGCAACAGCATCTGTTGCACTCAGCTGTGGGGCATCAACAGCTCCAGGC 1739
Qy 3020 AAGAACTCTGGCTGTGAAGAGATACCTAAAGAGTCAACAGCTCTCTGGGGATTTGGGGTTG 3079
Db 1740 AAGAACTCTGGCTGTGAAGAGATACCTAAAGAGTCAACAGCTCTCTGGGGATTTGGGGTTG 1799
Qy 3080 CTCGGAAGAACTCTTTGACACACTGCTGTGCTTGGAAATGCTGTTGAGTAATAATC 3139
Db 1800 CTCGGAAGAACTCTTTGACACACTGCTGTGCTTGGAAATGCTGTTGAGTAATAATC 1869
Qy 3140 TCTGGAACAGATTTGAAATACATGACCTGGATGAGCTGGAGACAGAAATTAACAAATTA 3199
|||||

Db 1860 TCTGGAACAGATTTGAAATACATGACCTGGATGAGCTGGAGACAGAAATTAACAAATTA 1919
Qy 3200 CACAGCTTAATTAACCTCTTAATTAAGAAATGCCAAACACAGAGAAAGAAATGACA 3259
Db 1920 CACAGCTTAATTAACCTCTTAATTAAGAAATGCCAAACACAGAGAAAGAAATGACA 1979
Qy 3260 AGAATTAATTAAGAAATTAAGTGGCAAGTTTGTGGAATTTGTTTAACTAACAAATTTG 3319
Db 1980 AGAATTAATTAAGAAATTAAGTGGCAAGTTTGTGGAATTTGTTTAACTAACAAATTTG 2039
Qy 3320 GCTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3379
Db 2040 GCTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2099
Qy 3380 TTTTGTCTGCTTCTTCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3439
Db 2100 TTTTGTCTGCTTCTTCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2159
Qy 3440 GACCCACCTCCCAATCCCGAGGGGACCCGACAGGCCGCGAAGAAATAGAAAGAGGTGG 3499
Db 2160 GACCCACCTCCCAATCCCGAGGGGACCCGACAGGCCGCGAAGAAATAGAAAGAGGTGG 2219
Qy 3500 AGAAGAGACAGAGACAGATCCATTCGATTAATTAATTAATTAATTAATTAATTAATTAAT 3559
Db 2220 AGAAGAGACAGAGACAGATCCATTCGATTAATTAATTAATTAATTAATTAATTAATTAAT 2279
Qy 3560 CGATCTGCGAGAGCTGTGCTCTTCACTACACACCGCTTGAAGAGACTTACTCTTATTTG 3619
Db 2280 CGATCTGCGAGAGCTGTGCTCTTCACTACACACCGCTTGAAGAGACTTACTCTTATTTG 2339
Qy 3620 AACGAGATTTGTGAACTTCTGGGAGCGAGGGGGTGGGAAGCCCTCAATTTTGTGGAA 3679
Db 2340 AACGAGATTTGTGAACTTCTGGGAGCGAGGGGGTGGGAAGCCCTCAATTTTGTGGAA 2399
Qy 3680 TCTCTCACTAATTTGAGTGAAGAACATTAAGTAATGCTTACTGCTCAATGGCAC 3739
Db 2400 TCTCTCACTAATTTGAGTGAAGAACATTAAGTAATGCTTACTGCTCAATGGCAC 2459
Qy 3740 AGCCATAGCAGTATGAGTGGAGAGAGATAGAGTTATAGATTAATTAACACACTTATAG 3799
Db 2460 AGCCATAGCAGTATGAGTGGAGAGAGATAGAGTTATAGATTAATTAACACACTTATAG 2519
Qy 3800 AGCTATTGCCACATACCTTAAGAAATTAAGACAGGGCTTGGAAAGATTTTGTCTATAGA 3859
Db 2520 AGCTATTGCCACATACCTTAAGAAATTAAGACAGGGCTTGGAAAGATTTTGTCTATAGA 2579
Qy 3860 TGGGTGGCAAGTGTCAAAAAGTAGTGTGATTTGGATGGCTGCTTAAGGGAAAGATGA 3919
Db 2580 TGGGTGGCAAGTGTCAAAAAGTAGTGTGATTTGGATGGCTGCTTAAGGGAAAGATGA 2639
Qy 3920 GACGAGCTGAGCCAGCAGAGATGGGTGGAGCAGATCTCGAGATCTAGA 3971
Db 2640 GACGAGCTGAGCCAGCAGAGATGGGTGGAGCAGATCTCGAGATCTAGA 2691

RESULT 12
US-09-827-688-4
: Sequence 4, Application US/09827688
: Publication No. US20030165476A1
: GENERAL INFORMATION:
: APPLICANT: ORSON, FRANK
: APPLICANT: KINSEY, BERNA
: APPLICANT: BHOGAL, BALBR
: TITLE OF INVENTION: MACROAGGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION
: FILE REFERENCE: P01949051/10004014
: CURRENT APPLICATION NUMBER: US/09/827,688
: PRIOR APPLICATION NUMBER: 60/195,680
: PRIOR FILING DATE: 2000-04-07
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: Patentin version 3.0
: SEQ ID NO 4
```


QY	2365	TCCTTTAAGCAATCCTCAGAGAGGGAGCCAGAAATTTGTAACGCACAGTTTAAATGTGGAC	2422
Db	1558	TCCTTTAAGCAATCCTCAGAGAGGGAGCCAGAAATTTGTAACGCACAGTTTAAATGTGGAC	1612
QY	2425	GGGAATTTTCTACTGTATATCAACACAACTGTTTAAATAGTACTGTGTTAAATGTACTT	2484
Db	1618	GGGAATTTTCTACTGTATATCAACACAACTGTTTAAATAGTACTGTGTTAAATGTACTT	1677
QY	2485	GGAGTACTGAAGGGTCAAAATPAACACTGAGAGAGTGCACACATGCACATCCTCCATGCAGAA	2544
Db	1678	GGAGTACTGAAGGGTCAAAATPAACACTGAGAGAGTGCACACATGCACATCCTCCATGCAGAA	1737
QY	2545	TAAACAATTTATTAACATGTGGCAGAGAAATGAGAAAAGCAATGATATGCCCCCTCCATCA	2604
Db	1738	TAAACAATTTATTAACATGTGGCAGAGAAATGAGAAAAGCAATGATATGCCCCCTCCATCA	1797
QY	2605	GTGAGCAAAATAGATGTTTCATCAATATATTCATGGGCTGCTATTTAAACAGAGTGGGTGA	2666
Db	1798	GTGAGCAAAATAGATGTTTCATCAATATATTCATGGGCTGCTATTTAAACAGAGTGGGTGA	1855
QY	2665	ATAACACAAATGGGTCCGAGATCTTCAGACCTGGAGAGAGCGATATGAGGACAAATTGGA	2724
Db	1858	ATAACACAAATGGGTCCGAGATCTTCAGACCTGGAGAGAGATATGAGGACAAATTGGA	1911
QY	2725	GAACTGAAATTTATTAATATTAATACTAGTAAATAATGAAACCATTTAGAGATACACCCACCA	2786
Db	1918	GAACTGAAATTTATTAATATTAATACTAGTAAATAATGAAACCATTTAGAGATACACCCACCA	1977
QY	2785	AGGCCAAAGAGAAAGTGGTGCAGAGAGAAAAAGCAGTGGGAAATAGAGCTTTGTGCC	2844
Db	1978	AGGCCAAAGAGAAAGTGGTGCAGAGAGAAAAAGCAGTGGGAAATAGAGCTTTGTGCC	2033
QY	2845	TTGGGTTCTTGGGAGCAGCAGAGAACTATGGGCTGCACGTCAATGACGCTGACGGTAC	2904
Db	2032	TTGGGTTCTTGGGAGCAGCAGAGAACTATGGGCTGCACGTCAATGACGCTGACGGTAC	2091
QY	2905	AGGCACAGACATTTTGTCTGTATATPAGAGCAGCAGCAGAACAAATTTGCTGAGGGCTATTG	2966
Db	2092	AGGCACAGACATTTTGTCTGTATATPAGAGCAGCAGCAGAACAAATTTGCTGAGGGCTATTG	2151
QY	2965	AGGGCAGCAACAGCATCTGTTCGACATCTACAGTGTGGGGCATCAACAGCTCCAGGCCAAGAA	3024
Db	2152	AGGGCAGCAACAGCATCTGTTCGACATCTACAGTGTGGGGCATCAACAGCTCCAGGCCAAGAA	2211
QY	3025	TCCTGGCTGTGGAAAGATACCTAAAGATCAACAAGCTCTGGGGATTTGGGGTTGCTCTG	3084
Db	2212	TCCTGGCTGTGGAAAGATACCTAAAGATCAACAAGCTCTGGGGATTTGGGGTTGCTCTG	2271
QY	3085	GAAACATCATTTGGACCACTGCTGTGCTTGGGATGCTATGTTGGAGTAAATATCTCTGG	3144
Db	2272	GAAACATCATTTGGACCACTGCTGTGCTTGGGATGCTATGTTGGAGTAAATATCTCTGG	2331
QY	3145	AACAGATTTTGGAAATPACATGACCTGAGTGGAGTGGAGAGAGAAATTAACATTPACACAA	3204
Db	2332	AACAGATTTTGGAAATPACATGACCTGAGTGGAGTGGAGAGAGAAATTAACATTPACACAA	2391
QY	3265	TATTTGGAATTTAGATTAATGGGCAAGTTTGTGGAATTTGGTTTAAACATTAACAAATTTGGCTGT	3324
Db	2452	TATTTGGAATTTAGATTAATGGGCAAGTTTGTGGAATTTGGTTTAAACATTAACAAATTTGGCTGT	2511
QY	3325	GGTATATTAATTAATTTATCATATATGATAGTGGAGGCTGTGTAGGTTTAAAGATAGTTTGTG	3384
Db	2512	GGTATATTAATTAATTTATCATATATGATAGTGGAGGCTGTGTAGGTTTAAAGATAGTTTGTG	2571
QY	3385	CTGTACTTTTCTATATAGTAATAGAGTTTGGCAGGGATATTCACATTAATCGTTTCAGACCC	3444
Db	2572	CTGTACTTTTCTGTATAGTAATAGAGTTTGGCAGGGATATTCACATTAATCGTTTCAGACCC	2631
QY	3445	ACCTTCCCAATCCGAGGGGAGCCGACAGAGGCCCGAAGAGATAGAAAGAGAGTGGAGAGA	3504

[illegible]

|||||
Db 598 ATGGGTACTGTTGGAGGAAGCAACCACTGATTTTGGCATCGATGCTTAAG 657
QY 1471 CATATGTTACAGAGTCAATATGTTGGGCCACACATGCTGTATGCCACAGACCCA 1350
Db 658 CATATGTTACAGAGTCAATATGTTGGGCCACACATGCTGTATGCCACAGACCCA 717
QY 1531 ACCCACAAGAGTAGTATTGTTAAATGTGACAGAAATTTTAAATGTCGAAAAATGACA 1590
Db 718 ACCCACAAGAGTAGTATTGTTAAATGTGACAGAAATTTTAAATGTCGAAAAATGACA 777
QY 1591 TGGTAGAACAGATGATGAGATATATCAGTTTATGGATCAAAGCCTTAAGCCATGTG 1650
Db 778 TGGTAGAACAGATGATGAGATATATCAGTTTATGGATCAAAGCCTTAAGCCATGTG 837
QY 1651 TAAATTTAACCCCACTGCTGTATAGTTTAAAGTCACTGATTTGAAATATGATCTATA 1710
Db 838 TAAATTTAACCCCACTGCTGTATAGTTTAAAGTCACTGATTTGAAATATGATCTATA 897
QY 1711 CCATAGTAGTAGCGGGAGATGATTAATGGAGAAAGAGATPAAAAACCTGCTTTCA 1770
Db 898 CCATAGTAGTAGCGGGAGATGATTAATGGAGAAAGAGATPAAAAACCTGCTTTCA 957
QY 1771 ATATCAGCACAAGCATATAGATATAGGTGCAGAAAGATATGCAATCTTTTATTAACCTTG 1830
Db 958 ATATCAGCACAAGCATATAGATATAGGTGCAGAAAGATATGCAATCTTTTATTAACCTTG 1017
QY 1831 ATATAGTAGCAATGATATA-----TACCAGCTATAGTGTGATPAGTTGTAACCCCTCAG 1884
Db 1018 ATATATATACCAATATGATATGATATCTACAGCTATACGTGACATGTTGTAACCCCTCAG 1077
QY 1885 TCATTTACACAGGCTGTCCAAAGGTATCCTTTGAGCCCAATTCCTATCATTTATGTGCC 1944
Db 1078 TCATTTACACAGGCTGTCCAAAGGTATCCTTTGAGCCCAATTCCTATCATTTATGTGCC 1137
QY 1945 CGGCTGGTTTGGGATTTCTAAATGTATATTAAGACGTTCAATGCAACAGGACCATGTA 2004
Db 1138 CGGCTGGTTTGGGATTTCTAAATGTATATTAAGACGTTCAATGCAACAGGACCATGTA 1197
QY 2005 CAATGTCAACAGTCAATGTATACATGGAATCAAGCCAGTATGATCACTCACTGTC 2064
Db 1198 CAATGTCAACAGTCAATGTATACATGGAATCAAGCCAGTATGATCACTCACTGTC 1257
QY 2065 TGTAAATGGCAGTCTAGCAGAAAGAGATGTAGTAATTTAGATCTGCCAATTTTCACAGACA 2124
Db 1258 TGTAAATGGCAGTCTAGCAGAAAGAGATGTAGTAATTTAGATCTGCCAATTTTCACAGACA 1317
QY 2125 ATGCTAAACCATTAATGTACAGTGAACACATCTGTAGAAATTTAATTTGTAACAAGCCA 2184
Db 1318 ATGCTAAACCATTAATGTACAGTGAACACATCTGTAGAAATTTAATTTGTAACAAGCCA 1377
QY 2185 ACAACATATCAAGAAAAAGTATCCGTATCCAGAGGGAGCCAGGAGAGCATTTGTTCAA 2244
Db 1378 ACAACATATCAAGAAAAAGTATCCGTATCCAGAGAGACCAAGAGAGCATTTGTTCAA 1437
QY 2245 TAGAAAAAATAGAAATATAGACAAGACCATTTGTAACATTTAGTAGACAATAGGATG 2304
Db 1438 TAGAAAAAATAGAAATATAGACAAGACCATTTGTAACATTTAGTAGAGCAAAATGGAATA 1497
QY 2305 CCCTTTAAACAGATGTAGTACGAATTAAGAAACAAATTTGAAAAATTAATTAATTAATTA 2364
Db 1498 ACCTTTAAACAGATGTAGTACGAATTAAGAAACAAATTTGAAAAATTAATTAATTAATTA 1557
QY 2365 TCTTTAAGCAATCTCTAGAGGGGAGCCAGAAATTTGTAAGCAGCAGTTTAAATTTGGAG 2424
Db 1558 TCTTTAAGCAATCTCTAGAGGGGAGCCAGAAATTTGTAAGCAGCAGTTTAAATTTGGAG 1617
QY 2425 GGGATTTTCTACTGTATTTCAACACACTGTTTAAATAGTACTGTTTAAATAGTACTT 2484
Db 1618 GGGATTTTCTACTGTATTTCAACACACTGTTTAAATAGTACTGTTTAAATAGTACTT 1677
QY 2485 GGAGTACTGAAGGTCAATTAACACTGAAGAGTGCACAATCACCTCCCATGCGAA 2544
|||||

Db 1678 GGAGTACTGAAGGTCAATTAACACTGAAGAGTGCACAATCACCTCCCATGCGAA 1737
QY 2545 TAAACAAATTTATTAACATGTGCGAGAAATAGGAAAGAAATGATATGCCCTCCATCA 2604
Db 1738 TAAACAAATTTATTAACATGTGCGAGAAATAGGAAAGAAATGATATGCCCTCCATCA 1797
QY 2605 GTGACAAATTTAGATGTTCAATATTTCTGCGCTGCTATTTAACAAAGAGATGTTGTA 2664
Db 1798 GTGACAAATTTAGATGTTCAATATTTACAGGCGCTGCTATTTAACAAAGAGATGTTGTA 1857
QY 2665 ATTAACAATTTGGTCCGAGATCTTCAGACCTTGAGAGAGCCATATAGAGGACAATTTGGA 2724
Db 1858 ATTAACAATTTGGTCCGAGATCTTCAGACCTTGAGAGAGGAGATATAGGACAATTTGGA 1917
QY 2725 GAATGAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2784
Db 1918 GAATGAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1977
QY 2785 AGCCAAAGAGAGATGTTCCAGAGAGAAAAAGAGCAGTGGGAATAGAGCTTTGTTCC 2844
Db 1978 AGCCAAAGAGAGATGTTCCAGAGAGAAAAAGAGCAGTGGGAATAGAGCTTTGTTCC 2031
QY 2845 TTGGGTTCTTTGGAGCAGCAGAGACATATGGGCTGCACGTCAATGACCTGACGCTAC 2904
Db 2032 TTGGGTTCTTTGGAGCAGCAGAGACATATGGGCTGCACGTCAATGACCTGACGCTAC 2091
QY 2905 AGGCCAGACATTTATTGCTGATATAGTGTGAGCAGCAGAACATTTTCTAGGGCTATTG 2964
Db 2092 AGGCCAGACATTTATTGCTGATATAGTGTGAGCAGCAGAACATTTTCTAGGGCTATTG 2151
QY 2965 AGGCCAGACATTTATTGCTGATATAGTGTGAGCAGCAGAACATTTTCTAGGGCTATTG 3024
Db 2152 AGGCCAGACATTTATTGCTGATATAGTGTGAGCAGCAGAACATTTTCTAGGGCTATTG 2211
QY 3025 TCTGCTGTGGGAAAGATTAACCTTAAGGATTAACACCTCTGGGGATTTGGGGTGTGCTG 3084
Db 2212 TCTGCTGTGGGAAAGATTAACCTTAAGGATTAACACCTCTGGGGATTTGGGGTGTGCTG 2271
QY 3085 GAAACATCATTTGTCACACACTGCTGTGCTGGAATGCTATGAGTAATTAATTTCTGCTG 3144
Db 2272 GAAACATCATTTGTCACACACTGCTGTGCTGGAATGCTATGAGTAATTAATTTCTGCTG 2331
QY 3145 AACAGATTTGGAATTAACATGACCTGATGAGAGTGGGACAGAGAAATTTAACATTTACAA 3204
Db 2332 AACAGATTTGGAATTAACATGACCTGATGAGAGTGGGACAGAGAAATTTAACATTTACAA 2391
QY 3205 GCTTAATACCTCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3264
Db 2392 GCTTAATACCTCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2451
QY 3265 TATTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3324
Db 2452 TATTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2511
QY 3325 GGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3384
Db 2512 GGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2571
QY 3385 CTGACTCTTATAGGAATTAAGATTAAGAGGATTTTACCATTAATGCTTTACAGCCC 3444
Db 2572 CTGACTCTTATAGGAATTAAGATTAAGAGGATTTTACCATTAATGCTTTACAGCCC 2631
QY 3445 ACCTCCCAATCCCGAGGGGAGCCGACAGGCCCAAGGAATTAAGAAAGAGTGGAGAGA 3504
Db 2632 ACCTCCCAATCCCGAGGGGAGCCGACAGGCCCAAGGAATTAAGAAAGAGTGGAGAGA 2691
QY 3505 GAGACAGACACATTCATTCATTTAGTGAAGGATCTTTAGCATTATCTGGAGAGATC 3564
Db 2692 GAGACAGACACATTCATTCATTTAGTGAAGGATCTTTAGCATTATCTGGAGAGATC 2751
QY 3565 TGGGAGGCTGTGCTCTTACGTACACAGGCTTGGAGAGCTTACCTTATGTTGTAACGA 3624
Db 2752 TGGGAGGCTGTGCTCTTACGTACACAGGCTTGGAGAGCTTACCTTATGTTGTAACGA 2811

Db 1326 CATCAGTGGACAAATTAGATGTTTCATCAAAATATTACAGGGCTGCTATTACAGAGATGG 1385
Qy 2660 TGTATATACAAATCATGGTCCGAGATCTTCAGACCTGGAGAGCGCATATAGAGGACAA 2719
Db 1386 TGTATATACAAATCATGGTCCGAGATCTTCAGACCTGGAGAGAGATATAGAGGACAA 1445
Qy 2720 TTGAGAGAAGTGAATTAATAAATTAAGTAGTAAAAATTGACCATTTAGGAGTAGCACC 2779
Db 1446 TTGGAGAAGTGAATTAATAAATTAAGTAGTAAAAATTGACCATTTAGGAGTAGCACC 1505
Qy 2780 CACCAAGGCAAAAGAGAGAGTGTGCAGAGAGAAAAAGACACATGGGAAATAGAGACTTT 2839
Db 1506 CACCAAGGCAAAAGAGAGAGTGTGCAGAGAGAAAAAGACACATGGGAAATAGAGACTTT 1565
Qy 2840 GTTCTTGGGTTCTTGGGAGCAGAGAGACATATGGGCTGCACGTCAATGACGCTGAC 2899
Db 1566 GTTCTTGGGTTCTTGGGAGCAGAGAGACATATGGGCTGCACGTCAATGACGCTGAC 1625
Qy 2900 GGTACAGGCGCAGACAAATTATTTGTGTATATAGTGCAGACAGACAAATTTGCTGAGGGC 2959
Db 1626 GGTACAGGCGCAGACAAATTATTTGTGTATATAGTGCAGACAGACAAATTTGCTGAGGGC 1685
Qy 2960 TATTGAGGCGCAAGACATCTGTGCACACTCAGACTCTGGGGCATCAACAGCTCCAGGC 3019
Db 1686 TATTGAGGCGCAAGACATCTGTGCACACTCAGACTCTGGGGCATCAACAGCTCCAGGC 1745
Qy 3020 AAGAACTCTGCTGTGAAAAGATACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTG 3079
Db 1746 AAGAACTCTGCTGTGAAAAGATACCTAAAGATCAACAGCTCTGGGGATTTGGGGTTG 1805
Qy 3080 CTCTGAAAACCTCATTTTGACACCACTGCTGTGCTTGGAAATGCTAGTGGAGTAAATAATC 3139
Db 1806 CTCTGAAAACCTCATTTTGACACCACTGCTGTGCTTGGAAATGCTAGTGGAGTAAATAATC 1865
Qy 3140 TCTGGACAGAGTTTGGAAATACATGACCTGATGAGTGGGACAGAGAAATTAACAATTA 3199
Db 1866 TCTGGACAGAGTTTGGAAATACATGACCTGATGAGTGGGACAGAGAAATTAACAATTA 1925
Qy 3200 CACAGCTTAATACACTCTTAATTAATGAAGATCCGAAACCCAGACAGAAAGAAATGAACA 3259
Db 1926 CACAGCTTAATACACTCTTAATTAATGAAGATCCGAAACCCAGACAGAAAGAAATGAACA 1985
Qy 3260 AGAATTAATGGAATTAAGTAAATGGCAAGTTTGTGAAATGTTTAAACATAACAATTG 3319
Db 1986 AGAATTAATGGAATTAAGTAAATGGCAAGTTTGTGAAATGTTTAAACATAACAATTG 2045
Qy 3320 GCTGTGATTAATTAATTAATTAATCAATGATGATGAGAGCTTGTAGGTTTAAGAAATAGT 3379
Db 2046 GCTGTGATTAATTAATTAATTAATCAATGATGATGAGAGCTTGTAGGTTTAAGAAATAGT 2105
Qy 3380 TTTTGTCTGTACTTTCTATAGTAAAGTAGAGTAGGAGAGATATTCACCATTAATGTTTCA 3439
Db 2106 TTTTGTCTGTACTTTCTATAGTAAAGTAGAGTAGGAGAGATATTCACCATTAATGTTTCA 2165
Qy 3440 GACCCACTCCCAATCCCGAGGGGACCCGACAGGCCCGAAGAAATAGAAAGAGAGTGG 3499
Db 2166 GACCCACTCCCAATCCCGAGGGGACCCGACAGGCCCGAAGAAATAGAAAGAGAGTGG 2225
Qy 3500 AGAGAGAGACAGACAGATCCATTCATTAAGTGAACGATCCTTAGACCTTAATCTGGGA 3559
Db 2226 AGAGAGAGACAGACAGATCCATTCATTAAGTGAACGATCCTTAGACCTTAATCTGGGA 2285
Qy 3560 CGATCTGGAGAGCTGTGCTCTCTCAGTACACCGCTTGAGAGACTTACTTGATTTGT 3619
Db 2286 CGATCTGGAGAGCTGTGCTCTCTCAGTACACCGCTTGAGAGACTTACTTGATTTGT 2345
Qy 3620 AACGAGATTTGTGAACCTTGTGGAGACGAGGGGTGGAGAACCTCAAAATTTGGTGA 3679
Db 2346 AACGAGATTTGTGAACCTTGTGGAGACGAGGGGTGGAGAACCTCAAAATTTGGTGA 2405
Qy 3680 TCTCTTACAGTATTTGAGAGTCAAGAACTTAAGAAATAGTCTTTTAACTTGTCAATGCCAC 3739
Db 2406 TCTCTTACAGTATTTGAGAGTCAAGAACTTAAGAAATAGTCTTTTAACTTGTCAATGCCAC 2465

Qy 3740 AGCCATAGCACTAGCTAGAGGACAGATAGGTTATAGAAATATTACAGACGCTATAG 3799
Db 2466 AGCTATAGCACTAGCTAGAGGACAGATAGGTTATAGAAATATTACAGAGCCTTATAG 2525
Qy 3800 AGCTATTTGCCACATTAACCTAGAGAAATPAAGACAGGCTTGGAAAGATTTTGTATA 3857
Db 2526 AGCTATTTGCCACATTAACCTAGAGAAATPAAGACAGGCTTGGAAAGATTTTGTATA 2583

Search completed: September 17, 2003, 22:52:22
Job time : 1376 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2003, 07:35:16 : Search time 11656 Seconds
(without alignments)
12988.382 Million cell updates/sec

Title: US-09-913-159A-10

Perfect score: 6229

Sequence: 1 ctgacgcgcctctgacgcgc.....attccccgaaagtgcac 6229

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrt1:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1055.8	16.9	1070	9	AJ281552
2	986.4	15.8	1013	12	BM438846
3	913.4	14.7	973	14	CD458281
4	902	14.5	1004	9	AJ281480

c	5	882.8	14.2	917	14	CD458286	CD458286 Fg08_09a0
c	6	882.8	14.2	956	9	AL959357	AL959357 AL959357
c	7	870.2	14.0	935	12	BG838279	BG838279 Gc01_10e0
c	8	862.6	13.8	918	14	CD459092	CD459092 Fg08_08e0
c	9	851.4	13.7	986	11	CNS08PE5	BX023239 Single re
c	10	841	13.5	841	9	AL042026	AL042026 DKFP434E
c	11	834.4	13.4	915	11	CNS09072	BX067467 Single re
c	12	832.8	13.4	872	14	CD459085	CD459085 Fg08_08d0
c	13	824.6	13.2	854	12	BM438950	BM438950 IPLV10049
c	14	822.4	13.2	893	9	AL879792	AL879792 AL879792
c	15	816.8	13.1	970	14	CD458333	CD458333 Fg08_09e0
c	16	814.4	13.0	1126	29	B2577702	B2577702 msh2_5533
c	17	807	12.8	863	11	CNS09LFC	BX063852 Single re
c	18	797.6	12.8	1011	29	B2576726	B2576726 msh2_5071
c	19	774.8	12.4	1336	29	B2575810	B2575810 msh2_4637
c	20	774.4	12.4	789	14	CD280920	CD280920 G44224.42
c	21	774.4	12.4	1574	29	B2572566	B2572566 msh2_2693
c	22	767.6	12.3	1370	29	B2571721	B2571721 msh2_2025
c	23	763.2	12.3	780	13	B0825693	B0825693 1030129B0
c	24	757	12.2	759	14	CD279661	CD279661 G43818_35
c	25	755.4	12.1	1463	29	B2571475	B2571475 msh2_1906
c	26	747.8	12.0	800	9	AJ281449	AJ281449 4A3A-P4D5
c	27	747.2	12.0	802	11	CNS08WC8	BX031332 Single re
c	28	743.6	11.9	832	12	BG923768	BG923768 602825893
c	29	742.6	11.9	786	14	CD458721	CD458721 Fg08_04e0
c	30	741	11.9	741	14	CD279174	CD279174 G44221_83
c	31	738.2	11.9	966	29	B2575002	B2575002 msh2_1255
c	32	737.8	11.8	1003	29	B2576686	B2576686 msh2_5053
c	33	736.8	11.8	950	29	B2571129	BX041089 Single re
c	34	733.8	11.8	782	11	CNS093V9	BX053933 Single re
c	35	733.4	11.8	817	11	CNS09DS1	BX0576702 msh2_5060
c	36	730.2	11.7	998	29	B2576702	CD281097 G44224.38
c	37	730	11.7	730	14	CD281097	B2571995 msh2_220
c	38	729.2	11.7	840	29	B2571995	CD279546 G44222.81
c	39	729	11.7	729	14	CD279546	B1687610 603312586
c	40	728.6	11.7	842	12	CD279322	CD279322 G44223.13
c	41	728	11.7	728	11	881	B2577222 msh2_5308
c	42	726.2	11.7	726	14	CD281811	CD281811 G44471.30
c	43	722.6	11.6	765	11	CNS093CH	BX040413 Single re
c	44	722.6	11.6	765	11	CNS093CH	BX040413 Single re
c	45	722	11.6	1394	29	B2580042	B2580042 msh2_908

ALIGNMENTS

RESULT 1
AJ281552
LOCUS 1070 bp mRNA linear EST 30-JUN-2000
DEFINITION 4A3A-P6F11-F Anopheles gambiae immune competent 4A3A-P6F11, mRNA sequence.

ACCESSION
AJ281552
VERSION
AJ281552.1 GI:6929432

KEYWORDS
SOURCE
ORGANISM

Anopheles gambiae (African malaria mosquito)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.

REFERENCE
1 (bases 1 to 1070)
Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansong, M., Soares, M. B.
and Kafatos, F. C.

Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines

Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)

JOURNAL
MEDLINE
PUBMED
20300950
10841361

COMMENT
CONTACT: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.

```

FEATURES
Source
Location/Qualifiers
1..1070
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P6F11"
/cell_line="Immune competent 4A3A"
/lab_host="E. coli DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1986) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT
263 a 283 c 255 g 269 t
ORIGIN

```

```

Query Match
Best Local Similarity 16.9%; Score 1055.8; DB 9; Length 1070;
Matches 1068; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 4471 TTGCTGGCGTTTTCATAGAGCTCGCCCGCCCTGACGACATCAAAAAATCGAGCTCA 4530
Db 1 TTGCTGGCGTTTTCATAGAGCTCGCCCGCCCTGACGACATCAAAAAATCGAGCTCA 60
QY 4531 AGTCAGAGGTGGCAAAACCCGACAGACTATTAAGATACAGGCGTTTCCCTCGGAAGC 4590
Db 61 AGTCAGAGGTGGCAAAACCCGACAGACTATTAAGATACAGGCGTTTCCCTCGGAAGC 120
QY 4591 TCCCTCGAGCTGCTCCGTTCACACCTGCGGCTTACCGGATACCTGCGCCCTTCTC 4650
Db 121 TCCCTCGAGCTGCTCCGTTCACACCTGCGGCTTACCGGATACCTGCGCCCTTCTC 180
QY 4651 CTTTCGGGAAGCGTGGCGCTTTCATATAGCTCAGCGTAGTATCAGTTCGGTGTAG 4710
Db 181 CTTTCGGGAAGCGTGGCGCTTTCATATAGCTCAGCGTAGTATCAGTTCGGTGTAG 240
QY 4711 GTCGTTGGCTCCAGCTGGGTGTGTGTCAGCAACCCCGCTTACCGCGCGTGGCC 4770
Db 241 GTCGTTGGCTCCAGCTGGGTGTGTGTCAGCAACCCCGCTTACCGCGCGTGGCC 300
QY 4771 TTATCCGGTAACTATCGTCTTGTAGTCCAAACCGGTAAAGACAGCTTATCGCCATGGCA 4830
Db 301 TTATCCGGTAACTATCGTCTTGTAGTCCAAACCGGTAAAGACAGCTTATCGCCATGGCA 360
QY 4831 GCAGCCACTGTAAACAGGATTAGCAGAGCGAGTATGTAGCGGTGTACAGAGTTCTTG 4890
Db 361 GCAGCCACTGTAAACAGGATTAGCAGAGCGAGTATGTAGCGGTGTACAGAGTTCTTG 420
QY 4891 AAGTGGGGCTACTACGCTGCTACACGTAAGAGACAGTATTTGATCTGGCTGTG 4950
Db 421 AAGTGGGGCTACTACGCTGCTACACGTAAGAGACAGTATTTGATCTGGCTGTG 480
QY 4951 AAGCAGTACCTTCGAAAAAGAGTTGTAGTCTTGTATCCGCAAAACCAACCGCT 5010
Db 481 AAGCAGTACCTTCGAAAAAGAGTTGTAGTCTTGTATCCGCAAAACCAACCGCT 540
QY 5011 GGTAGCGGTGTTTTTTTTTTGTTGCAAGCAGAGTTACGCGCAGAAAAAAGATCTCAA 5070
Db 541 GGTAGCGGTGTTTTTTTTTTGTTGCAAGCAGAGTTACGCGCAGAAAAAAGATCTCAA 600
QY 5071 GAAATGCTTTGATCTTTTACGGGGTGTGACCTGACGTAGGAGCAAAACTCACGTTAA 5130
Db 601 GAAATGCTTTGATCTTTTACGGGGTGTGACCTGACGTAGGAGCAAAACTCACGTTAA 660
QY 5131 GGAATTTTGGTCATGAGATTATCAAAAAGATCTTACCTAGATCTTTTAAATTAATAA 5190
Db 661 GGAATTTTGGTCATGAGATTATCAAAAAGATCTTACCTAGATCTTTTAAATTAATAA 720

```

```

QY 5191 TGAAGTTTAAATCATCTAAAGTATATAGTAAGTAACCTGTGCTGACGTTACCAATGC 5250
Db 721 TGAAGTTTAAATCATCTAAAGTATATAGTAAGTAACCTGTGACAGTTACCAATGC 780
QY 5251 TTATCAGTGAAGGACCTATCTCAGCGATCTGTCTATTTTGTATCATAGTTGCTGGA 5310
Db 781 TTATCAGTGAAGGACCTATCTCAGCGATCTGTCTATTTTGTATCATAGTTGCTGGA 840
QY 5311 CTCCCGGCGGTGATTAAGTATACATACGAGGAGGCTTACCATCTGGCCAGGTGCA 5370
Db 841 CTCCCGGCGGTGATTAAGTATACATACGAGGAGGCTTACCATCTGGCCAGGTGCA 900
QY 5371 ATGATACCGGACGACCCACGCTCACCGGCTCCAGATTATTCAGCAATAAACACGACCC 5430
Db 901 ATGATACCGGACGACCCACGCTCACCGGCTCCAGATTATTCAGCAATAAACACGACCC 960
QY 5431 GGAAGGCGGACGACGAGTGTCTGCAACTTATTCGCGCTTCATTCAGTCTATTAAT 5490
Db 961 GGAAGGCGGACGACGAGTGTCTGCAACTTATTCGCGCTTCATTCAGTCTATTAAT 1019
QY 5491 TGTTCGGGGAAGCTAGATAGTATGTTCCGCACTTATATGTTGGCCAAC 5541
Db 1020 TGTTCGGGGAAGCTAGATAGTATGTTCCGCACTTATATGTTGGCCAAC 1070

```

```

RESULT 2
BM438846/c 1013 bp mRNA linear EST 31-JAN-2002
LOCUS Iplv00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
DEFINITION Ictalurus punctatus (channel catfish)
ACCESSION BM438846
VERSION BM438846.1 GI:18460568
KEYWORDS EST.
SOURCE Ictalurus punctatus (channel catfish)
ORGANISM Ictalurus punctatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
REFERENCE 1 (bases 1 to 1013)
AUTHORS Feng J., Kucuktas H., Kocbas A., Li P. and Liu Z.
TITLE Transcription of channel catfish (Ictalurus punctatus): initial
analysis of expressed sequence tags from the liver
JOURNAL Unpublished
COMMENT Contact: Liu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@cesag.auburn.edu
Seq primer: M13 Reverse.

```

```

FEATURES
Source
Location/Qualifiers
1..1013
/organism="Ictalurus punctatus"
/mol_type="mRNA"
/db_xref="taxon:7998"
/clone_lib="Liver cDNA library"
/notes="Organ: Liver; Vector: pSport1; Site_1: NotI;
Site_2: SalI"
BASE COUNT
273 a 228 c 245 g 266 t 1 others
ORIGIN
Query Match
Best Local Similarity 15.8%; Score 986.4; DB 12; Length 1013;
Matches 1009; Conservative 1; Mismatches 2; Indels 2; Gaps 2;
QY 5044 ATTACGGCGCAAAAAAAGATCTCAAGAGATCCCTTGTATCTTTTACGGGGTCTGAC 5103
Db 1013 ATTACGGCGCAAAAAAAGATCTCAAGAGATCCCTTGTATCTTTTACGGGGTCTGAC 954
QY 5104 GCTCAGTGAACGAAACCTCAGCTTAAGGAGTTTGGTCATGAGATTATCAAAAAGATC 5163

```

```

Db      953 GCTGATGGAAGAACTCAAGTTAAGGATTTTGGTCATGAGATTTCAAAAAGATC 894
Qy      5164 TTCACGTGATCCTTTTAATTAATAAATGAAGTTTAATCAATCAATCAATATATATAG 5223
Db      893 TTCACGTGATCCTTTTAATTAATAAATGAAGTTTAATCAATCAATCAATATATATAG 834
Qy      5224 TAACTGTGTGACAGTTACCAATGCTTAATCAATGAGGACCTATCTCAGCGATCGT 5283
Db      833 TAACTGTGTGACAGTTACCAATGCTTAATCAATGAGGACCTATCTCAGCGATCGT 774
Qy      5284 CTATTCCTGTCATCATAGTTAGTCCCTGACTCCCGCTGCTGATAGTACATAGTACGAGAG 5343
Db      773 CTATTCCTGTCATCATAGTTAGTCCCGCTGACTCCCGCTGCTGATAGTACATAGTACGAGAG 714
Qy      5344 GCGTTACCATCTGGCCCCAGTGTGCAATGATPACCGGAGACCCAGCTCACCGGCTCCA 5403
Db      713 GCGTTACCATCTGGCCCCAGTGTGCAATGATPACCGGAGACCCAGCTCACCGGCTCCA 654
Qy      5404 GATTATACGCAATPAAACGACGCGGAGGCGGAGGCGGAGAGAGTGTCCGCACT 5463
Db      653 GATTATACGCAATPAAACGACGCGGAGGCGGAGGCGGAGAGAGTGTCCGCACT 594
Qy      5464 TTATCCGCTCCATCCAGTCTATTAATGTTGCCGGAAGCTAGATAGTATGTTCCGCA 5523
Db      593 TTATCCGCTCCATCCAGTCTATTAATGTTGCCGGAAGCTAGATAGTATGTTCCGCA 534
Qy      5524 GTTAATAGTTTGGCGACAGTGTGTCCTGCTACAGGCAATGCTGTGTCACGCTGTGCG 5583
Db      533 GTTAATAGTTTGGCGACAGTGTGTCCTGCTACAGGCAATGCTGTGTCACGCTGTGCG 474
Qy      5584 TTTGGTATGGCTCATTCAGTCCGGTCCCAAGATCAAGGCGAGTTACA-TGATGCC 5642
Db      473 TTTGGTATGGCTCATTCAGTCCGGTCCCAAGATCAAGGCGAGTTACA-TGATGCC 414
Qy      5643 CATGTGTGCAAAAAAGCGTTAGTCTCTGCTGCTCCGATCGTGTGTCAGAGTAAGTT 5702
Db      413 CATGTGTGCAAAAAAGCGTTAGTCTCTGCTGCTCCGATCGTGTGTCAGAGTAAGTT 355
Qy      5703 GCGCGCACTGTTATCATCATGTTATGACGACACTGCATTAATCTCTTACTGTGTCGCC 5762
Db      354 GCGCGCACTGTTATCATCATGTTATGACGACACTGCATTAATCTCTTACTGTGTCGCC 295
Qy      5763 ATCCGTAGATGCTTTCTGTGACGTGAGTACCAACCAAGTATCTAGAAATAGTG 5822
Db      294 ATCCGTAGATGCTTTCTGTGACGTGAGTACCAACCAAGTATCTAGAAATAGTG 235
Qy      5823 TATGCGGCGACGAGTGTCTGTGCGCGGCTCATATACGGAATATATACCGGCCCATAG 5882
Db      234 TATGCGGCGACGAGTGTCTGTGCGCGGCTCATATACGGAATATATACCGGCCCATAG 175
Qy      5883 CAGAACTTTAAAAAGTGTCTCATCTATTGAAAAAGTCTTCTGCGGCGGCAAACTCTCAAGAT 5942
Db      174 CAGAACTTTAAAAAGTGTCTCATCTATTGAAAAAGTCTTCTGCGGCGGCAAACTCTCAAGAT 115
Qy      5943 CTTAACCGCTGTTGAGATCCAGTGTGATTAACCCACTGTGTGACCACTGATCTTTCAGC 6002
Db      114 CTTAACCGCTGTTGAGATCCAGTGTGATTAACCCACTGTGTGACCACTGATCTTTCAGC 55
Qy      6003 ATCTTTACTTTTCAACAGCTTTTCTGGGTGAGCAAAAAACAGAAAGCAAAATGC 6056
Db      54 ATCTTTACTTTTCAACAGCTTTCTGGGTGAGCAAAAAACAGAAAGCAAAATGC 1

```

```

RESULT 3
CD458281/c 973 bp mRNA linear EST 03-JUN-2003
LOCUS Fg08_08h09_R Fg08_AAFC_ECORC_Fusarium_graminearum_complex_substrate
DEFINITION Gibberella zeae cDNA clone Fg08_08h09, mRNA sequence.
ACCESSION CD458281
VERSION CD458281.1 GI:31373021
KEYWORDS EST.
SOURCE Gibberella zeae

```

```

ORGANISM Gibberella zeae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE 1 (bases 1 to 973)
AUTHORS Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J., Hattori
,J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A., Spott,D. and
Tinker,N.A.
TITLE A cDNA library prepared from Fusarium graminearum grown on a
complex plant substrate
JOURNAL Unpublished
COMMENT Contact: Watson, Robert.J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1655
Fax: (613) 759-1701
Email: watsonrj@agr.gc.ca.
FEATURES
source location/Qualifiers
1. 973
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="taxon:5518"
/clone="Fg08_08h09"
/tissue_type="mycelium"
/dev_stage="asexual"
/lab_host="E. coli DH10B"
/clone_11b="Fg08_AAFC_ECORC_Fusarium_graminearum_complex_s
ubstrate"
/note="Vector: pBluescript II+; Site.1: EcoRI; Site.2:
XhoI; Fusarium graminearum grown on a complex plant
substrate-- wheat leaves treated to remove most of the low
molecular weight, water-soluble components."
BASE COUNT 264 a 215 c 238 g 251 t
ORIGIN
Query Match 14.7%; Score 913.4; DB 14; Length 973;
Best Local Similarity 98.7%; Pred. No. 9,7e-171;
Matches 925; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
Qy 5293 TCATCATAGTGTGCTGACTGCTCCCGCTGATAGTACATGATGAGGAGGCTTACCA 5352
Db 965 TCATCATAGTGTGCTGACTGCTCCCGCTGATAGTACATGATGAGGAGGCTTACCA 907
Qy 5353 TCTGCCCCAGTGTGCAATGATATACCGGAGACCCAGCTCACCGCTCCAGATTATCA 5412
Db 906 TCTGCCCCAGTGTGCAATGATATACCGGAGACCCAGCTCACCGCTCCAGATTATCA 847
Qy 5413 GCAATTAACCAACGACGCGGAGGCGGAGGCGGAGAGTGTCTGCAACTTTATCCGCC 5472
Db 846 GCAATTAACCAACGACGCGGAGGCGGAGGCGGAGAGTGTCTGCAACTTTATCCGCC 787
Qy 5473 TCCATCCAGTCTATTAATTTGTCGCGGAGCTAGTAAGTATGCGCAGTTATAGT 5532
Db 786 TCCATCCAGTCTATTAATTTGTCGCGGAGCTAGTAAGTATGCGCAGTTATAGT 727
Qy 5533 TTGCGCAAGTGTGTTGCACTTGTCTACAGCAATCGTGTGTCACGCTGTCTGTTGGTATG 5592
Db 726 TTGCGCAAGTGTGTTGCACTTGTCTACAGCAATCGTGTGTCACGCTGTCTGTTGGTATG 667
Qy 5593 GCTTATTCACGCTCGGTTCCCAAGATCAAGGCGAGTTACATGATCCCGCATGTTGTC 5652
Db 666 GCTTATTCACGCTCGGTTCCCAAGATCAAGGCGAGTTACATGATCCCGCATGTTGTC 607
Qy 5653 AAAAAAGCGTTAGTCTCTGCGTCCGATCGTGTGCAAGTAAGTGGCGCAGTG 5712
Db 606 AAAAAAGCGTTAGTCTCTGCGTCCGATCGTGTGCAAGTAAGTGGCGCAGTG 547
Qy 5713 TTATCATCATGTTATGAGCAGCACTGCATTAATCTCTTACTGTGTCATCCGTAAGA 5772
Db 546 TTATCATCATGTTATGAGCAGCACTGCATTAATCTCTTACTGTGTCATCCGTAAGA 487

```

FEATURES	source
JOURNAL MEDLINE PubMed	10841561
COMMENT	Contact: Dimopoulos G Fotis C. Kafatos laboratory European Molecular Biology Laboratory Meyerohofstrasse 1, 69117 Heidelberg, Germany. Location/Qualifiers 1. .1004
REFERENCE	1. .1004
AUTHORS	Dimopoulos, G., Casavant, T. L., Chang, S., Schieetz, T., Roberts, C., Donohue, M., Schultz, J. J., Benes, V., Bork, P., Ansoore, W., Soares, M. B., and Kafatos, F. C.
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
PROCEEDINGS	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
LOCUS	AJ281480
DEFINITION	4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION	AJ281480
VERSION	AJ281480.1 GI:5929360
KEYWORDS	EST.
SOURCE	Anopheles gambiae (African malaria mosquito)
ORGANISM	Anopheles gambiae Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prexygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
REFERENCE	1 (bases 1 to 1004)
AUTHORS	Dimopoulos, G., Casavant, T. L., Chang, S., Schieetz, T., Roberts, C., Donohue, M., Schultz, J. J., Benes, V., Bork, P., Ansoore, W., Soares, M. B., and Kafatos, F. C.
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
PROCEEDINGS	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
LOCUS	AJ281480
DEFINITION	4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION	AJ281480
VERSION	AJ281480.1 GI:5929360
KEYWORDS	EST.
SOURCE	Anopheles gambiae (African malaria mosquito)
ORGANISM	Anopheles gambiae Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prexygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
REFERENCE	1 (bases 1 to 1004)
AUTHORS	Dimopoulos, G., Casavant, T. L., Chang, S., Schieetz, T., Roberts, C., Donohue, M., Schultz, J. J., Benes, V., Bork, P., Ansoore, W., Soares, M. B., and Kafatos, F. C.
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
PROCEEDINGS	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
LOCUS	AJ281480
DEFINITION	4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION	AJ281480
VERSION	AJ281480.1 GI:5929360
KEYWORDS	EST.
SOURCE	Anopheles gambiae (African malaria mosquito)
ORGANISM	Anopheles gambiae Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prexygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
REFERENCE	1 (bases 1 to 1004)
AUTHORS	Dimopoulos, G., Casavant, T. L., Chang, S., Schieetz, T., Roberts, C., Donohue, M., Schultz, J. J., Benes, V., Bork, P., Ansoore, W., Soares, M. B., and Kafatos, F. C.
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
PROCEEDINGS	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
LOCUS	AJ281480
DEFINITION	4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION	AJ281480
VERSION	AJ281480.1 GI:5929360
KEYWORDS	EST.
SOURCE	Anopheles gambiae (African malaria mosquito)
ORGANISM	Anopheles gambiae Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prexygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
REFERENCE	1 (bases 1 to 1004)
AUTHORS	Dimopoulos, G., Casavant, T. L., Chang, S., Schieetz, T., Roberts, C., Donohue, M., Schultz, J. J., Benes, V., Bork, P., Ansoore, W., Soares, M. B., and Kafatos, F. C.
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
PROCEEDINGS	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
LOCUS	AJ281480
DEFINITION	4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION	AJ281480
VERSION	AJ281480.1 GI:5929360
KEYWORDS	EST.
SOURCE	Anopheles gambiae (African malaria mosquito)
ORGANISM	Anopheles gambiae Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prexygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
REFERENCE	1 (bases 1 to 1004)
AUTHORS	Dimopoulos, G., Casavant, T. L., Chang, S., Schieetz, T., Roberts, C., Donohue, M., Schultz, J. J., Benes, V., Bork, P., Ansoore, W., Soares, M. B., and Kafatos, F. C.
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
PROCEEDINGS	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
LOCUS	AJ281480
DEFINITION	4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION	AJ281480
VERSION	AJ281480.1 GI:5929360
KEYWORDS	EST.
SOURCE	Anopheles gambiae (African malaria mosquito)
ORGANISM	Anopheles gambiae Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prexygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
REFERENCE	1 (bases 1 to 1004)
AUTHORS	Dimopoulos, G., Casavant, T. L., Chang, S., Schieetz, T., Roberts, C., Donohue, M., Schultz, J. J., Benes, V., Bork, P., Ansoore, W., Soares, M. B., and Kafatos, F. C.
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
PROCEEDINGS	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
LOCUS	AJ281480
DEFINITION	4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION	AJ281480
VERSION	AJ281480.1 GI:5929360
KEYWORDS	EST.
SOURCE	Anopheles gambiae (African malaria mosquito)
ORGANISM	Anopheles gambiae Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prexygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
REFERENCE	1 (bases 1 to 1004)
AUTHORS	Dimopoulos, G., Casavant, T. L., Chang, S., Schieetz, T., Roberts, C., Donohue, M., Schultz, J. J., Benes, V., Bork, P., Ansoore, W., Soares, M. B., and Kafatos, F. C.
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
PROCEEDINGS	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
LOCUS	AJ281480
DEFINITION	4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION	AJ281480

BASE COUNT													cdna. The 4A3a is a directionally cloned and normalized cDNA library that was constructed from the 4A3a cell line o1dgo-T primed cDNA according to: Bonaldo, Lemon & Soares (1996) : Normalization and Subtraction: Two approaches To Facilitate Gene Discovery, Genome Research 6, 791-806."																																																																																																																																																																																																																																																																																																																																																									
ORIGIN																																																																																																																																																																																																																																																																																																																																																																						
252 a													262 c													244 g													244 t													2 others																																																																																																																																																																																																																																																																																																																		
Query Match													14.5%;													Score 902;													DB 9;													Length 1004;																																																																																																																																																																																																																																																																																																																		
Best Local Similarity													99.0%;													Pred. No. 1.8e-168;																																																																																																																																																																																																																																																																																																																																												
Matches 979;													Conservative													1;													Mismatches													2;													Indels													7;													Gaps													7;																																																																																																																																																																																																																																																														
QY	4464	GGCGCGCTGGTGGCGCTTTTTCATAGGCTCGCCCCCTGACAGCATCACAAAAATCG	4522	TTGAAAGCTCCCTGGGCGCTCTCCGTCGACCCCTGACGAGCTATTAAGATACAGGGCTTCCCCC	4583	83	ACGCTCAGTGTAGAGGTGGCGAAACCCGACAGGACTATTAAGATACAGGGCTTCCCCC	142	4584	TGGAAGCTCCCTGGCGCTCTCTGTTCCGACACCCTGCGCTTACCGGATACCTGTCCGC	4643	143	TGGAAGCTCCCTGGGCGCTCTCCGTCGACCCCTGCGCTTACCGGATACCTGTCCGC	202	4644	CTTCTCTCCCTTGGGGAACCGTGGCGCTTTCATATGCTCAGCGTGTAGTATCTCAGTTC	4703	203	CTTCTCTCCCTTGGGGAACCGTGGCGCTTTCATATGCTCAGCGTGTAGTATCTCAGTTC	262	4704	GGTGTAGTGTGCTGGCTCCAGAGCTGGGCTGTGTGGCAGCAACCCCGCTTACAGCCGACCG	4763	263	GGTGTAGTGTGCTGGCTCCAGAGCTGGGCTGTGTGGCAGCAACCCCGCTTACAGCCGACCG	322	4764	CTGGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAAACCCGGTAAAGACAGCTTATCGCC	4823	323	CTGGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAAACCCGGTAAAGACAGCTTATCGCC	382	4824	ACTGGCAGCACCACCTGTTAAACAGATTATGACAGAGCAGGATATGAGGCGGTGTACAGCA	4883	383	ACTGGCAGCACCACCTGTTAAACAGATTATGACAGAGCAGGATATGAGGCGGTGTACAGCA	442	4884	GTTCTTGAAGTGGGAGCCTTACTACGGCTACCTGATACCTGATGAGGAGCAGATTGTGATCTGGC	4943	443	GTTCTTGAAGTGGGAGCCTTACTACGGCTACCTGATACCTGATGAGGAGCAGATTGTGATCTGGC	502	4944	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	5003	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGCAGCATATACGCGCAGAAAAAAGG	5063	503	TCTGCTGAAGCCAGTTTACCTTCGAAAAAAGATTTGTAGTCTTGTATCCGGCAAAACAAC	562	5004	CACCGCTGTGTAGCGGTGTTTTTTTGTTCGACAGC

Db 920 TGCTCAATGATACCGCGAGA-CAACGGCTCA-CGGCTCCAGATT-TTACGCAATTAACCA 976
 QY 5424 GCCAGCCGGAAGCGCCGAGCGCAGAAAGT 5452
 Db 977 GCCAGCCGGAAG-GGCCGAGCGCAGAAAGT 1004
 RESULT 5
 CD458286/c 917 bp mRNA linear EST 03-JUN-2003
 LOCUS Fg08_09a03_R Fg08_AAFc_ECORC_Fusarium_graminearum_complex_substrate
 DEFINITION Gibberella zeae cDNA clone Fg08_09a03, mRNA sequence.
 ACCESSION CD458286
 VERSION CD458286.1 GI:31373026
 KEYWORDS EST.
 ORGANISM Gibberella zeae
 SOURCE Gibberella zeae
 ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Hypocryomycetidae; Hypocreales; Nectriaceae; Gibberella.
 REFERENCE 1 (bases 1 to 917)
 AUTHORS Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J., Hattori
 J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A., Sprott,D. and
 Tinker,N.A.
 TITLE A cDNA library prepared from Fusarium graminearum grown on a
 complex plant substrate
 JOURNAL Unpublished
 COMMENT Contact: Watson, Robert.J.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-food Canada
 Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
 CANADA
 Tel: (613) 759-1655
 Fax: (613) 759-1701
 Email: watsonrj@agr.gc.ca.
 FEATURES
 source 1..917
 Location/Qualifiers
 /organism="Gibberella zeae"
 /mol_type="mRNA"
 /strain="DAOM 180378"
 /db_xref="taxon:5518"
 /clone="Fg08_09a03"
 /tissue_type="Mycelium"
 /dev_stage="Asexual"
 /lab_host="E. coli DH10B"
 /clone_lib="Fg08_AAFc_ECORC_Fusarium_graminearum_complex_s
 ubstrate"
 /note="Vector: pBluescript II+; Site 1: EcoRI; Site 2:
 XhoI; Fusarium graminearum grown on a complex plant
 substrate--wheat leaves treated to remove most of the low
 molecular weight, water-soluble components."
 BASE COUNT 249 a 207 c 220 g 241 t
 ORIGIN
 Query Match 14.2% Score 882.8; DB 14; Length 917;
 Best Local Similarity 99.8%; Pred. No. 1,le-164;
 Matches 884; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5584 TTTGGTATGGCTTCATTCAGTCTCGGTTCCCAAGCATCAAGCGAGTATCATATCCCC 5643
 Db 677 TTTGGTATGGCTTCATTCAGTCTCGGTTCCCAAGCATCAAGCGAGTATCATATCCCC 618
 QY 5644 ATGTGTGCAAAAAAGCGGTAGCTCTCGGTTCCGATGTTGTCAAGATAGTTG 5703
 Db 617 ATGTGTGCAAAAAAGCGGTAGCTCTCGGTTCCGATGTTGTCAAGATAGTTG 558
 QY 5704 GCCGAGTGTATATCATATGTTATGCGAGCATGCAATATTCCTTACTGTATGCCA 5763
 Db 557 GCCGAGTGTATATCATATGTTATGCGAGCATGCAATATTCCTTACTGTATGCCA 498
 QY 5764 TCCGTAAGATGTTTCTGTGATCTGTGATGATCTCAACCAATCTTGTGAATATGTT 5823
 Db 497 TCCGTAAGATGTTTCTGTGATCTGTGATGATCTCAACCAATCTTGTGAATATGTT 438
 QY 5824 ATGCGGCGACGAGTTGCTTGGCGGCGGATACGAGATATACCGGACATATGAC 5883
 Db 437 ATGCGGCGACGAGTTGCTTGGCGGCGGATACGAGATATACCGGACATATGAC 378
 QY 5884 AGAAGCTTAAAGTCTCATTCATTTGAAACGTTCTGGGGCGAAACCTCAAGATC 5943
 Db 377 AGAAGCTTAAAGTCTCATTCATTTGAAACGTTCTGGGGCGAAACCTCAAGATC 318
 QY 5944 TTACCGCTGTAGATTCAGTTCATGATGATGATGATGATGATGATGATGATGATG 6003
 Db 317 TTACCGCTGTAGATTCAGTTCATGATGATGATGATGATGATGATGATGATGATG 258
 QY 6004 TCTTTTACTTTCACGAGGTTCTGTTGTTGAGCAAAACGAAAGCAAAATCCGCAAA 6063
 Db 257 TCTTTTACTTTCACGAGGTTCTGTTGTTGAGCAAAACGAAAGCAAAATCCGCAAA 198
 QY 6064 AAGGCAATTAAGGCGCACACGGAATGTAATCTCATCTCTCTCTCTCTCTCTCTCT 6123
 Db 197 AAGGCAATTAAGGCGCACACGGAATGTAATCTCATCTCTCTCTCTCTCTCTCTCT 138
 QY 6124 TGAACATTTATCAGGCTTATTTCTCATGACGATCATATTTTGAATGTTAGAAA 6183
 Db 137 TGAACATTTATCAGGCTTATTTCTCATGACGATCATATTTTGAATGTTAGAAA 78
 QY 6184 AATAACAATATGAGGTTCCGCGCACATTTCCCGCAAAAGTCCAC 6229
 Db 77 AATAACAATATGAGGTTCCGCGCACATTTCCCGCAAAAGTCCAC 32
 RESULT 6
 AL959357 956 bp mRNA linear EST 27-NOV-2002
 LOCUS AL959357 XGC-gastrulia Silurana tropicalis cDNA clone Tgas127X19 5',
 DEFINITION mRNA sequence.
 ACCESSION AL959357
 VERSION AL959357
 KEYWORDS EST.
 SOURCE AL959357.1 GI:25782952
 ORGANISM Silurana tropicalis (western clawed frog)
 Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Silurana.
 REFERENCE 1 (bases 1 to 956)
 AUTHORS Taylor R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.
 Sanger Xenopus tropicalis EST project 2002
 TITLE Taylor R.
 JOURNAL Unpublished
 COMMENT Contact: Taylor R
 Sanger Centre
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: tropesanger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: Tgas127X19.plk5b6
 Sequencing primer: SP6
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 FEATURES
 Location/Qualifiers

```
source
1. .956
/organism="Silurena tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="Tgas127k19"
/dev_stage="gastrula (stages 10.5-13 mixed)"
/lab_host="Escherichia coli xtl-blue"
/clone_lib="XGC-gastrula"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from stages
10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
into pCS107 with EcoRI at the 5' end and NotI at the 3'
end."
BASE COUNT      242 a      242 c      235 g      236 t      1 others
ORIGIN
Query Match      14.2% Score 882.8; DB 9; Length 956;
Best Local Similarity 99.7% Pred. No. 1.2e-164;
Matches 884; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
4464 GCGCGCGTGTGCGGCTTTTCCATAGAGCTCCGCCCTGAGCAGATCACAATAATCG 4523
|||||
65 GCGCGCGTGTGCGGCTTTTCCATAGAGCTCCGCCCTGAGCAGATCACAATAATCG 124
4524 AGCTCAAGTCAGAGGTGCGAACCAGACAGACTATAAGATACAGCGCTTCCCGC 4583
|||||
125 AGCTCAAGTCAGAGGTGCGAACCAGACAGACTATAAGATACAGCGCTTCCCGC 184
4584 TGAAGCTCCCTCGGCGCTCTCGTTCGACCGCTGCGGCTTACCGGATACGTCGCC 4643
|||||
185 TGAAGCTCCCTCGGCGCTCTCGTTCGACCGCTGCGGCTTACCGGATACGTCGCC 244
4644 CTTCCTCCCTCGGGAAGCGTGGCGCTTTCATAGCTCAGCTGTAGTATCTCAGTTC 4703
|||||
245 CTTCCTCCCTCGGGAAGCGTGGCGCTTTCATAGCTCAGCTGTAGTATCTCAGTTC 304
4704 GGTGTAGTCTGTGCTTCCAAAGCTGGGCTGTGTGACAGACCCCGCTTACGCCGACG 4763
|||||
305 GGTGTAGTCTGTGCTTCCAAAGCTGGGCTGTGTGACAGACCCCGCTTACGCCGACG 364
4764 CTGGGCGCTTATCGGTAACCTACGCTTGAAGTCCCAACCGGTAAGACACGACTATGCC 4823
|||||
365 CTGGGCGCTTATCGGTAACCTACGCTTGAAGTCCCAACCGGTAAGACACGACTATGCC 424
4824 ACTGGACAGACCACTGCTAAGATAGGACAGACGAGTATAGGCGGTCTACAGA 4883
|||||
425 ACTGGACAGACCACTGCTAAGATAGGACAGACGAGTATAGGCGGTCTACAGA 484
4884 GTTCTTGAAGTGTGGCTTAACCTACGCTTACAGTAAAGACAGTATTTGGTATCTGCCG 4943
|||||
485 GTTCTTGAAGTGTGGCTTAACCTACGCTTACAGTAAAGACAGTATTTGGTATCTGCCG 544
4944 TCTGCTGAAGCACTTACCTCGGAAAAAGTGTGCTTGAATCCGCGCAAAAC 5003
|||||
545 TCTGCTGAAGCACTTACCTCGGAAAAAGTGTGCTTGAATCCGCGCAAAAC 604
5004 CACCGCTGTAGCGGTGTTTGTGTTTGAACAGAGATTAAGCGCACAATAAAAGG 5063
|||||
605 CACCGCTGTAGCGGTGTTTGTGTTTGAACAGAGATTAAGCGCACAATAAAAGG 664
5064 ATCTCAAGAGATCTTTTGTATCTTTCTACGGGGTCTGACGCTCAGTGAAGCAAAATC 5123
|||||
665 ATCTCAAGAGATCTTTTGTATCTTTCTACGGGGTCTGACGCTCAGTGAAGCAAAATC 724
5124 ACGTTAAGGATTTTGGTATGAGATTAACAATAAGATCTTACCTTCTTTTAA 5183
|||||
725 ACGTTAAGGATTTTGGTATGAGATTAACAATAAGATCTTACCTTCTTTTAA 784
5184 TTAATAAAGATTTTAATCAATCTAAGATATATAGTAAACTGGTGCAGCTTA 5243
|||||
785 TTAATAAAGATTTTAATCAATCTAAGATATATAGTAAACTGGTGCAGCTTA 844
5244 CCAATGCTTAATCAAGTGAAGCACTATCTCAGCATCTGTATTTGCTTCAATCAATAGT 5303
```

```
|||||
Db 845 CCAATGCTTAATCAAGTGAAGCACTATCTCAGCATCTGTATTTGCTTCAATCAATAG 904
QY 5304 TGCCCTGACTCCCGCTGTGATGATACATACATACGAGGAGGCTTAC 5350
|||||
Db 905 TGCCCTGACTCCCGCTGTGATGATACATACATACGAGGAGGCTTAC 951
RESULT 7
Bg838279/c
LOCUS Bg838279 935 bp mRNA linear EST 25-MAY-2001
DEFINITION Gc01_10e07_R Gc01_AAFc_ECORC_cold_stressed_glycine_clandestina
ACCESSION Bg838279
VERSION Bg838279
KEYWORDS Bg838279.1 GI:14204601
SOURCE EST.
ORGANISM Glycine clandestina
Glycine clandestina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 935)
AUTHORS Singh,J.A., Farah,S., Chapados,J., Couroux,P., De Moors,A., Harris
,L.J., Hattori,J.I., Onelle,T., Robert,L.S., Sprolt,D. and Tinker
,N.A.
TITLE Expressed Sequence Tags from Cold-Stressed Glycine clandestina
Seedlings
JOURNAL Unpublished
COMMENT Contact: Singh,J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA
0C6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@em.agr.ca.
FEATURES
source
1. .935
/location/Qualifiers
/organism="Glycine clandestina"
/mol_type="mRNA"
/cultivar="1035"
/db_xref="taxon:45687"
/clone="Gc01_10e07"
/rissue_type="Leaves, stem"
/clone_lib="Gc01_AAFc_ECORC_cold_stressed_glycine_clandest
ina"
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;
Site 2: XhoI; Plants incubated at 2 degrees under 12 hours
of light/day. Harvested after only 2-3 days of cold
treatment. cDNA was prepared with the Uni-Zap cDNA kit
from Stratagene. Eco RI adapters were linked followed by
digest with Xho I/Eco RI and ligated to pBluescript."
BASE COUNT      243 a      213 c      227 g      238 t      14 others
ORIGIN
Query Match      14.0% Score 870.2; DB 12; Length 935;
Best Local Similarity 98.1% Pred. No. 3.6e-162;
Matches 874; Conservative 14; Mismatches 2; Indels 1; Gaps 1;
5340 GAGAGGCTTACCATCTGGCCCGCACTGCTGCAATGATACCGGAGACCCACGCTACCGGC 5399
|||||
928 GGGGGGCTTACCATCTGGCCCGCAKCGTGCATATATACCGGAGMCCACGCTCMCGCG 869
5400 TCACGATTTATCAGCAATTAACACAGCACGCCGGAAGGCGGACGCGCAGAAAGTGTCTGC 5459
|||||
868 YCCAGATTTATCAGCAATTAACACAGCACGCCGGAAGGCGGACGCGCAGAAAGTGTCTGC 809
5460 AACCTTATCCGCCCTCCATCCAGCTATTAATTTGTCGGGGAAGCTGAGTAAGTACTCTC 5519
|||||
808 AACTTTATCCGCCCTCCATCCAGCTATTAATTTGTCGGGGAAGCTGAGTAAGTACTCTC 749
5520 GCCAGTTAATAGTTTGGCAACGTTTGGCATTTGTAACAGCATCTGTGTCACGCTC 5579
```



```

|||||
Db 748 GCCAGTATATAGTTGGCCCAAGTTGTGCCATTGCTACAGGCATCGTGGTCCGCTC 689
QY 5580 GTCTGTTGGTATGGCTTATTCAGCTCCGGTTCCCAACGATCAGCGAGTTACATGATC 5639
Db 688 GTCTGTTGGTATGGCTTATTCAGCTCCGGTTCCCAACGATCAGCGAGTTACATGATC 629
QY 5640 CCCCATGTTGGCAAAAAAGGGGTAGTCCCTGGCTCCGATCGTGTGCAGAGTAA 5699
Db 628 CCCCATGTTGGCAAAAAAGGGGTAGTCCCTGGCTCCGATCGTGTGCAGAGTAA 569
QY 5700 GTT-GGGCCAGTGTATACATCATGTTATGGCAGCACTGCATATTCCTCTACTGTCA 5758
Db 568 GTTGGCCGCGAGTGTATACATCATGTTATGGCAGCACTGCATATTCCTCTACTGTCA 509
QY 5759 TGCATCCGTAGATGCTTTCTGTGACTGTGTGACTACTCAACCAAGTCAATCTGAGAT 5818
Db 508 TGCCATCCGTAGATGCTTTCTGTGACTGTGTGACTACTCAACCAAGTCAATCTGAGAT 449
QY 5819 AGTGTATGGCGGACCGAGTGTCTGTGCTGGCGCTCAATACGGATATACCGGCGAC 5878
Db 448 AGTGTATGGCGGACCGAGTGTCTGTGCTGGCGCTCAATACGGATATACCGGCGAC 389
QY 5879 ATAGCAGAACTTTAAAGTGCATCATTTGAAAAAGCTTTCTGGGGGAAAACTGTCAA 5938
Db 388 ATAGCAGAACTTTAAAGTGCATCATTTGAAAAAGCTTTCTGGGGGAAAACTGTCAA 329
QY 5939 GGATCTTACCGCTGTGTGATCCAGTGTGATGTAACCCAGTCGTGACCCCAACTGATCT 5998
Db 328 GGATCTTACCGCTGTGTGATCCAGTGTGATGTAACCCAGTCGTGACCCCAACTGATCT 269
QY 5999 CAGATCTTTACTTTCACACGCGTTCTGGGTGGCAAAAAAGGAGGCAAAATGGCG 6058
Db 268 CAGATCTTTACTTTCACACGCGTTCTGGGTGGCAAAAAAGGAGGCAAAATGGCG 209
QY 6059 CAAAAAGGATAAGGCGACACGGAATGTGATATCTCATACTCTCTTTTCAT 6118
Db 208 CAAAAAGGATAAGGCGACACGGAATGTGATATCTCATACTCTCTTTTCAT 149
QY 6119 ATTATGAAGCATTTATAGGGTATTTGTCTCATGAGCGGATACATATTTGAATGATTT 6178
Db 148 ATTATGAAGCATTTATAGGGTATTTGTCTCATGAGCGGATACATATTTGAATGATTT 89
QY 6179 AGAAAAATTAACAATAGGGGTCCGCGCATATTCGCCGAAGATGGCAC 6229
Db 88 AGAAAAATTAACAATAGGGGTCCGCGCATATTCGCCGAAGATGGCAC 38

RESULT 8
CD459092/c 918 bp mRNA linear EST 03-JUN-2003
LOCUS Fg08_08e02_R Fg08_AAFC_ECORC_Fusarium_graminearum_complex_substrate
DEFINITION Gibberella zeae cDNA clone Fg08_08e02, mRNA sequence.
ACCESSION CD459092
VERSION CD459092.1 GI:31373832
KEYWORDS EST.
SOURCE Gibberella zeae
ORGANISM Gibberella zeae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE 1 (bases 1 to 918)
AUTHORS Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J., Hattori
J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A., Sprott,D. and
Tinker,N.A.
TITLE A cDNA library prepared from Fusarium graminearum grown on a
complex plant substrate
JOURNAL Unpublished
COMMENT Contact: Watson, Robert.J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, KIA 0C6,
CANADA
Tel.: (613) 759-1655

```

```

FEATURES
    source
        Location/Qualifiers
            1..918
                /organism="Gibberella zeae"
                /mol_type="mRNA"
                /strain="DAOM 180378"
                /db_xref="taxon:5518"
                /clone="Fg08_08e02"
                /ltsue_type="Mycelium"
                /dev_stage="Asexual"
                /lab_host="E. coli DH108"
                /clone_lib="Fg08_AAFC_ECORC_Fusarium_graminearum_complex-s
                ubstrate"
                /note="Vector: pBluescript II+; Site_1: EcoRI; Site_2:
                XhoI; Fusarium graminearum grown on a complex plant
                substrate-- wheat leaves treated to remove most of the low
                molecular weight, water-soluble components."
BASE COUNT      247 a      208 c      221 g      239 t      3 others
ORIGIN
Query Match      13.8%; Score 862.6; DB 14; Length 918;
Best Local Similarity 99.4%; Pred. No. 1.2e-160;
Matches 884; Conservative 2; Mismatches 1; Indels 2; Gaps 2;

5343 GGGCTTACATCTGGCCCCCAGTGTGCAATGA-TACCGCGAGACCCAGCTCAGCGCTC 5401
|||||
Db 918 GGGCTTACATCTGGCCCCCAGTGTGCAATGA-TTCCCGAGMCCAGCTCAGCGCTC 859
5402 CAGATTATTCAGCAATTAACCCAGCGCGGAAGGGCGAGAGTGGTCTGCA 5461
|||||
Db 858 CAGATTATTCAGCAATTAACCCAGCGCGGAAGGGCGAGAGTGGTCTGCA 799
5462 CTTTATCCGCTCCATCCAGTCTATTAATTGTTGGCGGGAAGCTAGAGTAAGTTCGC 5521
|||||
Db 798 CTTTATCCGCTCCATCCAGTCTATTAATTGTTGGCGGGAAGCTAGAGTAAGTTCGC 739
5522 CAGTTAATAGTTGGGCAACGTTGTGCAATGCTACAGCACTCGTGTGCAACCTCGT 5581
|||||
Db 738 CAGTTAATAGTTGGGCAACGTTGTGCAATGCTACAGCACTCGTGTGCAACCTCGT 679
5582 CGTTTGTATGCTTCATTCAGCTCCGGTCCCAAGATCAAGGAGTTCATGATCCC 5641
|||||
Db 678 CGTTTGTATGCTTCATTCAGCTCCGGTCCCAAGATCAAGGAGTTCATGATCCC 619
5642 CCATGTTGTGCAAAAAAGCGGTTAGCTCTTCCTGCTCCGATCGTGTGCAAGTAAGT 5701
|||||
Db 618 CCATGTTGTGCAAAAAAGCGGTTAGCTCTTCCTGCTCCGATCGTGTGCAAGTAAGT 559
5702 TGGCCGAGTGTATACATCA-TGGTTATGGCAGACAGCAATTCCTTACGTCTCAG 5760
|||||
Db 558 TGGCCGAGTGTATACATCA-TGGTTATGGCAGACAGCAATTCCTTACGTCTCAG 499
5761 CCATCCGTAAAGTCTTCTGTGACTGTGAGTACTCAACCAAGTCAATTCGAGAAATAG 5820
|||||
Db 498 CCATCCGTAAAGTCTTCTGTGACTGTGAGTACTCAACCAAGTCAATTCGAGAAATAG 439
5821 TGTATGGCGGACGAGTGTGCTCTTGCCCGGCGTAAATACGGGATTAATCCGCGCACAT 5880
|||||
Db 438 TGTATGGCGGACGAGTGTGCTCTTGCCCGGCGTAAATACGGGATTAATCCGCGCACAT 379
5881 AGCAGAACTTTAAAGTCTCATCTTTGAAAGGCTTCTGGGGCGAAACCTCAAG 5940
|||||
Db 378 AGCAGAACTTTAAAGTCTCATCTTTGAAAGGCTTCTGGGGCGAAACCTCAAG 319
5941 ATCTTACCGCTGTGTGAGATCCAGTTCGATGTAACCACTGTGACCCCAACTGATCTTCA 6000
|||||
Db 318 ATCTTACCGCTGTGTGAGATCCAGTTCGATGTAACCACTGTGACCCCAACTGATCTTCA 259
6001 GCATCTTTTACTTTCACAGCGTTTCTGGGTGAGCAAAAGAGGAGCAAAATGCGCGCA 6060
|||||
Db 258 GCATCTTTTACTTTCACAGCGTCTCTGTGAGCAAAAGAGGAGCAAAATGCGCGCA 199

```

QY 6061 AAAAAGGAATAGGCGGACACGGAATGTTGAATCTCATCTCTCTTTTCAATAT 6120
 |||||
 Db 198 AAAAGGAATAGGCGGACACGGAATGTTGAATCTCATCTCTCTTTTCAATAT 139
 QY 6121 TATGAGCATTTTACAGGGTATGTCTCATGCGGATTCATATTGAATGATTTAG 6180
 |||||
 Db 138 TATGAGCATTTTACAGGGTATGTCTCATGCGGATTCATATTGAATGATTTAG 79
 QY 6181 AAAAATAACAAATAGGGGTTCCGCGACATTTCCGGAAGTGCAC 6229
 |||||
 Db 78 AAAAATAACAAATAGGGGTTCCGCGACATTTCCGGAAGTGCAC 30
 RESULT 9
 CDS08PES/c 986 bp mRNA linear HTC 07-JAN-2003
 LOCUS Single read from an extremity of a full-length cDNA clone made from
 DEFINITION Anopheles gambiae total adult females. 3-PRIME end of clone
 FK0AA34AH01 of strain 6-9 of Anopheles gambiae (African malaria
 mosquito).
 ACCESSION BX022329
 VERSION BX022329
 KEYWORDS HTC.
 SOURCE Anopheles gambiae (African malaria mosquito)
 ORGANISM Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoides;
 Anopheles.
 REFERENCE 1 (bases 1 to 986)
 Genoscope.
 AUTHORS Direct Submission
 TITLE Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 FEATURES
 source
 1. 986
 /organism="Anopheles gambiae"
 /mol_type="mRNA"
 /strain="6-9"
 /db_xref="taxon:7165"
 /clone="FK0AA34AH01"
 /plasmid="pME18S-FL"
 /note="end : 3-PRIME"
 BASE COUNT 264 a 233 c 237 g 252 t
 ORIGIN
 Query Match 13.7%: Score 851.4; DB 11; Length 986;
 Best Local Similarity 97.4%: Pred. No. 2e-158;
 Matches 889; Conservative 0; Mismatches 16; Indels 8; Gaps 2;
 QY 5325 GATACTACGATACGAGGAGGCTTACCATCTGCGCCAGTGCATATGATACCGGAGA 5384
 |||||
 Db 986 GATACTACGATACGAGGAGGCTTACCATCTGCGCCAGTGCATATGATACCGGAGA 927
 QY 5385 CCCACGCTCACCGGCTCCAGATTTATACGAATTAACACGACCGGGAAGGCCGAGCG 5444
 |||||
 Db 926 CCCACGCTCACCGGCTCCAGATTTATACGAATTAACACGACCGGGAAGGCCGAGCG 867
 QY 5445 CAGAAGTGTGCTCCAGCTTATTCGCCCTCATCCACTCTATTAATTTGCCGGGAAGC 5504
 |||||
 Db 866 CAGAAGTGTGCTCCAGCTTATTCGCCCTCATCCACTCTATTAATTTGCCGGGAAGC 807
 QY 5505 TAGAGTAAGTAGTTCGCGCAGTAAATTAATTTGCCGCAAGTGTGCTTACAGCAT 5564
 |||||
 Db 806 TAGAGTAAGTAGTTCGCGCAGTAAATTAATTTGCCGCAAGTGTGCTTACAGCAT 747
 QY 5565 CGTGGTTCACGCTGCTGTTGGTATAGCTTCATTCAGCTCCGGTCCCAACGATCAAG 5624
 |||||
 Db 746 CGTGGTTCACGCTGCTGTTGGTATAGCTTCATTCAGCTCCGGTCCCAACGATCAAG 687
 QY 5625 GCGAGTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTGGTCTCCGAT 5684
 |||||
 Db 686 GCGAGTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTGGTCTCCGAT 627

QY 5685 CGTTGTACAGAAATAGTTGGCGCAGTGTATATCATCTGATGTTATGCGACACTGCATAA 5744
 |||||
 Db 626 CGTTGTCTGTTTTTGTGTGGCGCAGTGTATATCATCTGATGTTATGAAATTAATGCAATA 567
 QY 5745 TTCTCTTACTGTCTATGCCATTCGTAAGATGCTTTTGTGTGACTGTGATCTACACCA 5804
 |||||
 Db 566 TTCTCTTACTGTCTATGCCATTCGTAAGATGCTTTTGTGTGACTGTGATCTACACCA 507
 QY 5805 GTCATTCGAGAAATAGTATAGCGCGCAGTGTGCTTGGCCGCGTCAATACGGA 5864
 |||||
 Db 506 GTCATTCGAGAAATAGTATAGCGCGCAGTGTGCTTGGCCGCGTCAATACGGA 447
 QY 5865 TAATACCGCGCCACATAGCAGAACTTAAAGTCTCATCATTTGAAAGCGTTCTCGG 5924
 |||||
 Db 446 TAATACCGCGCCACATAGCAGAACTTAAAGTCTCATCATTTGAAAGCGTTCTCGG 387
 QY 5925 GCGAAACCTCTCAAGATCTTACCGCTGTGTGACATCCAGTTGATACCACTCGTC 5984
 |||||
 Db 386 GCGAAACCTCTCAAGATCTTACCGCTGTGTGACATCCAGTTGATACCACTCGTC 327
 QY 5985 ACCCACTGATCTTACAGATCTTTTACTTTCACAGGCTTCTGGTGAGCAAAACAGG 6044
 |||||
 Db 326 ACCCACTGATCTTACAGATCTTTTACTTTCACAGGCTTCTGGTGAGCAAAACAGG 267
 QY 6045 AAGCCAAATGCGCCCAAAAAAGGAAATAGGCGCAGCAGCAATGTTGAA-----TAC 6097
 |||||
 Db 266 AAGCCAAATGCGCCCAAAAAAGGAAATAGGCGCAGCAGCAATGTTGAA-----TAC 207
 QY 6098 TCATACTCTTC-CTTTTCAATATTTATGAAGCATTTATACAGGTTATTTGTCTCATAGC 6156
 |||||
 Db 206 GAGTACTCTTCCTGCTTTTCAATATTTATGAAGCATTTATACAGGTTATTTGTCTCATAGC 147
 QY 6157 GGATACATATTTGAATGATTTAGAAATTAACAAATPAGGGTTCGCGCACATTTCC 6216
 |||||
 Db 146 GGATACATATTTGAATGATTTATGAATTAACAAATPAGGGTTCGCGCACATTTCC 87
 QY 6217 CGAAAGTGCAC 6229
 |||||
 Db 86 CGAAAGTGCAC 74
 RESULT 10
 AL042026/c 841 bp mRNA linear EST 29-FEB-2000
 LOCUS DKFP434E111.r1 434 (synonym: hies3) Homo sapiens cDNA clone
 DEFINITION DKFP434E111 5', mRNA sequence.
 ACCESSION AL042026
 VERSION AL042026.1 GI:5421372
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 841)
 AUTHORS Poustka A., Klein M., Mewes H.W., Gassenhuber J. and Wiemann S.
 TITLE EST (Poustka, et al.)
 JOURNAL Unpublished
 COMMENT
 Contact: Poustka A.J.
 Department Lehrach
 Max-Planck-Institute for Molecular Genetics
 Inestrasse 73, 14195 Berlin, Germany
 Tel: +49-30-84131623
 Fax: +49-30-84131128
 Email: poustka@mpg-berlin-dahlem.mpg.de
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by DKFZ (German Cancer Research Center,
 Heidelberg/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 No sl sequence available.
 This clone (DKFP434E111) is available at the RZPD in Berlin.

Please contact the R2PD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@r2pd.de.

FEATURES

Location/Qualifiers
1. 841

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434E111"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="434 (synonym: hles3)"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"

BASE COUNT 226 a 192 c 202 g 221 t

ORIGIN

Query Match 13.5%; Score 841; DB 9; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.3e-156;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

5355 TGGCCCGCAGTGTGCAATGATATACGCGGAGACCCGACCGGCTCCAGATTATTCACG 5414
|||||
841 TGGCCCGCAGTGTGCAATGATATACGCGGAGACCCGACCGGCTCCAGATTATTCACG 782
5415 AATAAACGACGACGCGGAGAGGCGCGAGCGAGAGTGTCTGCACTTATCCGCCCTC 5474
|||||
781 AATAAACGACGACGCGGAGAGGCGCGAGCGAGAGTGTCTGCACTTATCCGCCCTC 722
5475 CATCCAGTCTATTAATTGTGCGCGGAGAGTGAAGTAAGTGTCCAGTTAATAGTTT 5534
|||||
721 CATCCAGTCTATTAATTGTGCGCGGAGAGTGAAGTAAGTGTCCAGTTAATAGTTT 662
5535 GCGGACGCTGTGTCATGCTGCTACAGCATGCTGTGTCACGCTGCTGTGTTGGATAGC 5594
|||||
661 GCGGACGCTGTGTCATGCTGCTACAGCATGCTGTGTCACGCTGCTGTGTTGGATAGC 602
5595 TTCATTCAGTCCGCTGCCAAGCATGAGCGAGTATGATGATCCCATGTTGTGCA 5654
|||||
601 TTCATTCAGTCCGCTGCCAAGCATGAGCGAGTATGATGATCCCATGTTGTGCA 542
5655 AAAAGCGGTAGCTCTCGTCTCGATGCTGTGTCAGAAAGTAAGTTGGCCGAGTGT 5714
|||||
541 AAAAGCGGTAGCTCTCGTCTCGATGCTGTGTCAGAAAGTAAGTTGGCCGAGTGT 482
5715 ATCACTCATGTTATGAGACATGCAATATCTTCTGTCATGCAATGCTGTTAAGT 5774
|||||
481 ATCACTCATGTTATGAGACATGCAATATCTTCTGTCATGCAATGCTGTTAAGT 422
5775 CTTTCTGTGACTGTGAGTACTCAACCAAGTCAATTCATGAGTAATGATGATGCGGACG 5834
|||||
421 CTTTCTGTGACTGTGAGTACTCAACCAAGTCAATTCATGAGTAATGATGATGCGGACG 362
5835 GAGTTGCTCTGCGCGCGGTCAATACGGGATTAATCCGCGCATACAGACTTTAAA 5894
|||||
361 GAGTTGCTCTGCGCGCGGTCAATACGGGATTAATCCGCGCATACAGACTTTAAA 302
5895 AGTGTCTCATGTTGAAAGCTCTGCGGCGGAGAAATCTCAACAGACTTAAACCCCTTT 5954
|||||
301 AGTGTCTCATGTTGAAAGCTCTGCGGCGGAGAAATCTCAACAGACTTAAACCCCTTT 242
5955 GAGATCCAGTTCGATGTAACCACTGTCACCCCAAGTCAATTCATGAGATCTTTTACTTT 6014
|||||
241 GAGATCCAGTTCGATGTAACCACTGTCACCCCAAGTCAATTCATGAGATCTTTTACTTT 182
6015 CACGAGCGTTTCTGGGTGAGCAAAAAAGAGGCAAAATGCCGCAAAAAAGGAATAG 6074
|||||
181 CACGAGCGTTTCTGGGTGAGCAAAAAAGAGGCAAAATGCCGCAAAAAAGGAATAG 122
6075 GCGGACGAGAAATGTTCAATACATCTCTCTCTTTTCAATTAATTGAGCATTTA 6134
|||||
121 GCGGACGAGAAATGTTCAATACATCTCTCTCTTTTCAATTAATTGAGCATTTA 62
6135 TCAGGTTATTGTCTCATGAGCGGATACATATTGTAATTTAGAAAAATTAACAAT 6194

Db 61 TCAGGTTATTGTCTCATGAGCGGATACATATTGTAATTTAGAAAAATTAACAAT 2

QY 6195 A 6195

Db 1 A 1

RESULT 11

CNS09072/c

LOCUS

DEFINITION

Single read from an extremity of a full-length cDNA clone made from
Anopheles gambiae total adult females. 3-PRIME end of clone
FK0AC51AF10 of strain 6-9 of Anopheles gambiae (African malaria
mosquito).

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage ;
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

Location/Qualifiers

1. 915

/organism="Anopheles gambiae"

/mol_type="mRNA"

/strain="6-9"

/db_xref="taxon:7165"

/clone="FK0AC51AF10"

/plasmid="PME18S-PL"

/note="end : 3-PRIME"

BASE COUNT 241 a 224 c 218 g 232 t

ORIGIN

Query Match 13.4%; Score 834.4; DB 11; Length 915;
Best Local Similarity 99.3%; Pred. No. 4.7e-155;
Matches 838; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

5386 CCAGCTCACGCGCTCCAGATTATTCAGCAATTAACACGACCGGAGGCGGACGCGC 5445
|||||
914 CCGCGCGCGGCGCGCCAGATTATTCAGCAATTAACACGACCGGAGGCGGACGCGC 855
5446 AGAAGTGTGCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAGCT 5505
|||||
854 AGAAGTGTGCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAGCT 795
5506 AGAGTAAGTATGTCGCAAGTATAGTTTGGCAAGTGTGTTGCCATTGCTACAGGATC 5565
|||||
794 AGAGTAAGTATGTCGCAAGTATAGTTTGGCAAGTGTGTTGCCATTGCTACAGGATC 735
5566 GTGGTTCACGCTGCTGTTGGTATGCTTCATTCACGCTCCGTTCCACAGATCAAG 5625
|||||
734 GTGGTTCACGCTGCTGTTGGTATGCTTCATTCACGCTCCGTTCCACAGATCAAG 675
5626 CGAGTTACATGATCCCATGTTGTCGCAAAAAAGGGTTAGTCTGCTGCTCCGATC 5685
|||||
674 CGAGTTACATGATCCCATGTTGTCGCAAAAAAGGGTTAGTCTGCTGCTCCGATC 615
5686 GTTTCAGAAAGTAAGTTGGCGCAGTGTATTCATCATGATGATGAGACACTGCATAT 5745
|||||
614 GTTTCAGAAAGTAAGTTGGCGCAGTGTATTCATCATGATGATGAGACACTGCATAT 555
5746 TCTCTTACTGTCATCCATCCGTAAGATGCTTTCTGTGACGTGTGAGTACTCAACCAAG 5805
|||||
554 TCTCTTACTGTCATCCATCCGTAAGATGCTTTCTGTGACGTGTGAGTACTCAACCAAG 495

QY 5806 TCATCTGAGAAATAGTGTATGCGCGACCGAGTTGCTCTTCCCGCGCTCAATACGGGAT 5865
 |||||||
 Db 494 TCATCTGAGAAATAGTGTATGCGCGACCGAGTTGCTCTTCCCGCGCTCAATACGGGAT 435
 |||||||
 QY 5866 AATACCGCGCCACATACAGAACTTTAAAGTGTCTCATCTGAGAAAACGTTCTTGGGG 5925
 |||||||
 Db 434 AATACCGCGCCACATACAGAACTTTAAAGTGTCTCATCTGAGAAAACGTTCTTGGGG 375
 |||||||
 QY 5926 CGAAACCTCAAGGATCTTACCGCTGTGATCCGTTCCGATTAACCAACGCTGCGCA 5985
 |||||||
 Db 374 CGAAACCTCAAGGATCTTACCGCTGTGATCCGTTCCGATTAACCAACGCTGCGCA 315
 |||||||
 QY 5986 CCCAAGTATCTTACGATCTTCTTACTTCCACGACCTTCTGGGTGAGCAAAAACAGGA 6045
 |||||||
 Db 314 CCCAAGTATCTTACGATCTTCTTACTTCCACGACCTTCTGGGTGAGCAAAAACAGGA 255
 |||||||
 QY 6046 AGCGAAATGCGCGCAAAAAGGAAATAGGCGCGACGAGAAATGTGAATCTCATCTC 6105
 |||||||
 Db 254 AGCGAAATGCGCGCAAAAAGGAAATAGGCGCGACGAGAAATGTGAATCTCATCTC 195
 |||||||
 QY 6106 TTCTTTTCAATATTTATTAAGCATTTATCAGGTTATGTCTCATAGCGGATACATA 6165
 |||||||
 Db 194 TTCTTTTCAATATTTATTAAGCATTTATCAGGTTATGTCTCATAGCGGATACATA 135
 |||||||
 QY 6166 TTGAATGTATTTAGAAAATTAACAATAGGCGCTCCGCGACATTTCCCGAAAAGTG 6225
 |||||||
 Db 134 TTGAATGTATTTAGAAAATTAACAATAGGCGCTCCGCGACATTTCCCGAAAAGTG 75
 |||||||
 QY 6226 CCAC 6229
 |||||
 Db 74 CCAC 71

RESULT 12

CD459085/c 872 bp mRNA linear EST 03-JUN-2003
 LOCUS Fg08_08d03_R Fg08_AAFc_ECORC_Fusarium_graminearum_complex_substrate
 DEFINITION Giberella zeae cDNA clone Fg08_08d03, mRNA sequence.

ACCESSION CD459085
 VERSION CD459085.1 GI:31373825

KEYWORDS EST.
 SOURCE Giberella zeae

ORGANISM

Giberella zeae
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Hypocreomycetidae; Hypocreales; Nectriaceae; Giberella.

REFERENCE 1 (bases 1 to 872)
 AUTHORS Watson,R.J., Heys,R., Chapados,J., Couroux,P., Harris,L.J., Hattori

J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A., Sprott,D. and
 Tinker,N.A.

TITLE A cDNA library prepared from Fusarium graminearum grown on a
 complex plant substrate

JOURNAL Unpublished
 COMMENT Contact: Watson, Robert.J.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-Food Canada
 Bldg. 20, Central Experimental Farm, Ottawa, Ontario, KIA 0C6,
 CANADA

REL: (613) 759-1655
 Fax: (613) 759-1701
 Email: watsonrj@agr.gc.ca.

Location/Qualifiers

1. 872

FEATURES
 SOURCE /organism="Giberella zeae"
 /mol_type="mRNA"
 /strain="DAOM 180378"
 /db_xref="taxon:5518"
 /clone="Fg08_08d03"
 /issue_type="Mycelium"
 /dev_stage="Asexual"
 /lab_host="E. coli DH10B"
 /clone_id="Fg08_AAFc_ECORC_Fusarium_graminearum_complex_s
 ubstrate"

/note="Vector: pBluescript II+; Site_1: EcoRI; Site_2:

BASE COUNT 235 a 194 c 210 g 228 t 5 others
 ORIGIN
 Query Match 13.4%; Score 832.8; DB 14; Length 872;
 Best Local Similarity 99.2%; Pred. No. 9,7e-155;
 Matches 831; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 5385 CCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACCAACCGCACCGGAAAGCGCGAG 5444
 |||||||
 Db 872 CCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACCAACCGCACCGGAAAGCGCGAG 813
 |||||||
 QY 5445 CAGAAGTGTCTCGCACTTATTCGGCTCCATTCATCTATTAATTTGTTCCGGGAAGC 5504
 |||||||
 Db 812 CAGAAGTGTCTCGCACTTATTCGGCTCCATTCATCTATTAATTTGTTCCGGGAAGC 753
 |||||||
 QY 5505 TAGAGTAAGTAGTTCGCGAGTTAATAGTTTGGCGAACGTTGGCCATTGCTACAGGCAT 5564
 |||||||
 Db 752 TAGAGTAAGTAGTTCGCGAGTTAATAGTTTGGCGAACGTTGGCCATTGCTACAGGCAT 693
 |||||||
 QY 5565 CGTGTGTACAGCTCGTCTGTTGGTATGCTTCAATTCAGCTCCGTTCCCAACGATCAAG 5624
 |||||||
 Db 692 CGTGTGTACAGCTCGTCTGTTGGTATGCTTCAATTCAGCTCCGTTCCCAACGATCAAG 633
 |||||||
 QY 5625 CGCAGTTACATGATCCCGCATGTTGTGCAAAAAAGCGGTTAAGTCTCTTCCGCTCCGAT 5684
 |||||||
 Db 632 CGCAGTTACATGATCCCGCATGTTGTGCAAAAAAGCGGTTAAGTCTCTTCCGCTCCGAT 573
 |||||||
 QY 5685 CGTTGTCAAGTAAGTTGGCCGCGAGTTCATCATCATGTTATGCGACACATCATTA 5744
 |||||||
 Db 572 CGTTGTCAAGTAAGTTGGCCGCGAGTTCATCATCATGTTATGCGACACATCATTA 513
 |||||||
 QY 5745 TTCTCTTACTGTCTATGCCATCCGTAAGATGCTTTCTGTACATGCTGATCTCAACCA 5804
 |||||||
 Db 512 TTCTCTTACTGTCTATGCCATCCGTAAGATGCTTTCTGTACATGCTGATCTCAACCA 453
 |||||||
 QY 5805 GTCATCTGAGAAATAGTGTATGCGGCGACGAGTTCGTTGCCCGCGCAATACGGGA 5864
 |||||||
 Db 452 GTCATCTGAGAAATAGTGTATGCGGCGACGAGTTCGTTGCCCGCGCAATACGGGA 393
 |||||||
 QY 5865 TAATACCGCGCCACATAGCAAACTTTAAAGTCTCATCATGTTGAAAGAGTTCTTCGGG 5924
 |||||||
 Db 392 TAATACCGCGCCACATAGCAAACTTTAAAGTCTCATCATGTTGAAAGAGTTCTTCGGG 333
 |||||||
 QY 5925 GCGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCCATTAACCCACTGCTGC 5984
 |||||||
 Db 332 GCGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCCATTAACCCACTGCTGC 273
 |||||||
 QY 5985 ACCCACTGATCTTACGATCTTACTTCCACAGGCTTTCGGGTGAGCAAAAACAGG 6044
 |||||||
 Db 272 ACCCACTGATCTTACGATCTTACTTCCACAGGCTTTCGGGTGAGCAAAAACAGG 213
 |||||||
 QY 6045 AAGCAAAATGCGCGCAAAAAGGGAATAGGCGGACACGGAATGTTGAATCTCATACT 6104
 |||||||
 Db 212 AAGCAAAATGCGCGCAAAAAGGGAATAGGCGGACACGGAATGTTGAATCTCATACT 153
 |||||||
 QY 6105 CTTCCTTTTCAATATTTATTGAAGCATTTATCAGGCTTATTTGCTCATGAGCGGATCAT 6164
 |||||||
 Db 152 CTTCCTTTTCAATATTTATTGAAGCATTTATCAGGCTTATTTGCTCATGAGCGGATCAT 93
 |||||||
 QY 6165 ATTGAATGTATTTAGAAAATTAACAATAGGCGTCCGCGACATTTCCCGGAAA 6222
 |||||||
 Db 92 ATTGAATGTATTTAGAAAATTAACAATAGGCGTCCGCGACATTTCCCGGAAA 35
 |||||||

RESULT 13

BM438950/c 854 bp mRNA linear EST 31-JAN-2002
 LOCUS BM438950
 DEFINITION IDLVr00491 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
 sequence.

ACCESSION BM438950

VERSION BM438950.1 GI:18460672
 KEYWORDS EST.
 SOURCE Ictalurus punctatus (channel catfish)
 ORGANISM Ictalurus punctatus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
 Ictaluridae; Ictalurus.
 1 (bases 1 to 854)
 REFERENCE Peng, J., Kucuktas, H., Kocbas, A., Li, P. and Liu, Z.
 Transcription of channel catfish (Ictalurus punctatus): initial
 analysis of expressed sequence tags from the liver
 Unpublished
 JOURNAL Contact: Liu ZJ
 The Fish Molecular Genetics and Biotechnology Laboratory,
 Department of Fisheries and Allied Aquacultures and Program of Cell
 and Molecular Biosciences
 Auburn University
 203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
 Tel: 334 844 4054
 Fax: 334 844 9208
 Email: zliu@acesag.auburn.edu
 COMMENT Seq primer: M13 Reverse
 Location/Qualifiers
 1..854
 /organism="Ictalurus punctatus"
 /mol_type="mRNA"
 /db_xref="taxon:7998"
 /clone_lib="Liver cDNA library"
 /note="Organ: Liver; Vector: pSport1; Site_1: NotI;
 Site_2: SalI"
 BASE COUNT 226 a 198 c 221 g 209 t
 ORIGIN
 Query Match 13.2%; Score 824.6; DB 12; Length 854;
 Best Local Similarity 99.3%; Pred. No. 4.1e-153;
 Matches 849; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
 OY 5203 TCATCTAAGATATATATAGTAACTGGTGTGACAGTATACCAATGCTTATCAAGTGG 5262
 DB 854 TCAATCTAAGATATATAGTAACTGGTGTGACAGTATACCAATGCTTATCAAGTGG 795
 OY 5263 GCACCTATCTCAGCATCTGTCTATTTGCTTCATCCATAGTTCCTGATCCCGCTGCTG 5322
 DB 794 GCACCTATCTCAGCATCTGTCTATTTGCTTCATCCATAGTTCCTGATCCCGCTGCTG 735
 OY 5323 TAGATACTACATACGGAGGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGCA 5382
 DB 734 TAGATACTACATACGGAGGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGCA 675
 OY 5383 GACCCAGCTCAGCGCTCAGATTATCAGCAATAAACGACGCGGAAGGGCGAG 5442
 DB 674 GACCCAGCTCAGCGCTCAGATTATCAGCAATAAACGACGCGGAAGGGCGAG 615
 OY 5443 CGCAGAAGTGTCTGCAACTTATCCGCTCCATCCAGTCTATTAAATGTTGCCGGAA 5502
 DB 614 CGCAGAAGTGTCTGCAACTTATCCGCTCCATCCAGTCTATTAAATGTTGCCGGAA 555
 OY 5503 GCTAAGTAAGTATGTTGCCAGTATTAAGTTTGGCAACGTTGGCCATTGCTACAGGC 5562
 DB 554 GCTAAGTAAGTATGTTGCCAGTATTAAGTTTGGCAACGTTGGCCATTGCTACAGGC 495
 OY 5563 ATCGGGGTGTCAGCTCGCTGTTGGTATAGCTTATTCAGTCCGGTTCGCAAGATGA 5622
 DB 494 ATCGGGGTGTCAGCTCGCTGTTGGTATAGCTTATTCAGTCCGGTTCGCAAGATGA 435
 OY 5623 AGCGAGTTACATGATCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCGGTCTCGG 5682
 DB 434 AGCGAGTTACATGATCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCGGTCTCGG 376
 OY 5683 ATCGTTGTGCAAGTAAGTGGCCGAGTGTATCATCTCATGTTATGGCAGACATGAT 5742
 DB 375 ATCGTTGTGCAAGTAAGTGGCCGAGTGTATCATCTCATGTTATGGCAGACATGAT 316

OY 5743 AATTCCTTACTGTGATGCCATCCGATAGATGCTTTTGTGATGCTGTGAGTACTCAACC 5802
 DB 315 AATTCCTTACTGTGATGCCATCCGATAGATGCTTTTGTGATGCTGTGAGTACTCAACC 256
 OY 5803 AAGTCATTCGTGAAATAGTATGTCGGGACGACGAGTGTGCTTCCCGGCTCAATACGG 5862
 DB 255 AAGTCATTCGTGAAATAGTATGTCGGGACGACGAGTGTGCTTCCCGGCTCAATACGG 196
 OY 5863 GATTAATACGGCGCCACATATGCAAGACTTTAAAGTCTATATATGGAAGAGTCTTTCG 5922
 DB 195 GATTAATACGGCGCCACATATGCAAGACTTTAAAGTCTATATATGGAAGAGTCTTTCG 136
 OY 5923 GGGCGAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGAA-CCGACTCG 5981
 DB 135 GGGCGAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGAAAGCCACTCG 76
 OY 5982 TGCACCAACTGATCTTACGATCTTTTACTTTACACAGCGTTTCTGGTGAGCAAAAC 6041
 DB 75 TGCACCAACTGATCTTACGATCTTTTACTTTACACAGCGATCTGGTGAGCAAAAC 16
 OY 6042 AGGAAGCCAAATGCG 6056
 DB 15 AGGAAGCCAAATGCG 1
 RESULT 14
 AL879792 993 bp mRNA linear EST 15-SEP-2002
 LOCUS AL879792 XCC-egg Silurana tropicalis cDNA clone Tegg019c18 3', mRNA
 DEFINITION sequence.
 ACCESSION AL879792
 VERSION AL879792
 SOURCE AL879792.1 GI:22900057
 ORGANISM EST.
 Silurana tropicalis (western clawed frog)
 Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Silurana.
 1 (bases 1 to 993)
 REFERENCE Taylor, R., Ashurst, J.L., Cronling, M.D.R., Zorn, A.M. and Rogers, J.
 Authors Taylor, R., Ashurst, J.L., Cronling, M.D.R., Zorn, A.M. and Rogers, J.
 Title Sanger Xenopus tropicalis EST project 2002
 Journal Unpublished
 Comment Sanger Centre
 Contact: Taylor R
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: Tegg019c18.q1kT7
 Sequencing primer: 17
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 FEATURES
 source
 1..993
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="Tegg019c18"
 /dev_stage="egg"
 /lab_host="Escherichia coli XL1-blue"
 /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
 was oligo dt primed from 5ug of poly A+ RNA from egg.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end"
 BASE COUNT 275 a 218 c 219 g 278 t
 ORIGIN
 Query Match 13.2%; Score 822.4; DB 9; Length 993;
 Best Local Similarity 99.5%; Pred. No. 1.1e-152;
 Matches 823; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 OY 5403 AGATTATCAGCAATAAACAGCCGCGAGGGCGCGCAGAAAGTGTCTGCAAC 5462
 DB 5403 AGATTATCAGCAATAAACAGCCGCGAGGGCGCGCAGAAAGTGTCTGCAAC 5462

Db 993 AGATTATACCAATTAACACGAGCCGGAAGGCGGAGCGACGAAGTGTCTGTCATC 934
 Oy 5463 TTATATCCGCTCCATCCAGTCTATTAATTTGTCGGGAGAGCTAGATAGTATGCC 5522
 Db 933 TTATATCCGCTCCATCCAGTCTATTAATTTGTCGGGAGAGCTAGATAGTATGCC 874
 Oy 5523 AGTTAATAGTTGGGCAAGTGTGTCATGCTACAGCAAGTGTGTCAGCTGCTC 5582
 Db 873 AGTTAATAGTTGGGCAAGTGTGTCATGCTACAGCAAGTGTGTCAGCTGCTC 814
 Oy 5583 GTTTGGTATGCTTCATTCACCTCCGCTTCCCAAGATCAAGGCGAGTTACATGATCCC 5642
 Db 813 GTCTGTATGCTTCATTCACCTCCGCTTCCCAAGATCAAGGCGAGTTACATGATCCC 754
 Oy 5643 CATGTGTGCAAAAAAGGCGGTAGCTCTGCTGCTCCGATCGTGTGCAAGATAT 5702
 Db 753 CATGTGTGCAAAAAAGGCGGTAGCTCTGCTGCTCCGATCGTGTGCAAGATAT 694
 Oy 5703 GCGCGCAGTGTATACATCATGATGTTATGACAGACATGATATCTCTTACTGTCATCC 5762
 Db 693 GCGCGCAGTGTATACATCATGATGTTATGACAGACATGATATCTCTTACTGTCATCC 634
 Oy 5763 ATCCGTAGATGCTTTTCTGTGACTGTGATGATCTCAACCAAGTATTTCTGAGATATG 5822
 Db 633 ATCCGTAGATGCTTTTCTGTGACTGTGATGATCTCAACCAAGTATTTCTGAGATATG 574
 Oy 5823 TATCGCGGACCGAGTGTCTCTGCTCCGCGCTCAATAGGATATATCCGCGCATAG 5882
 Db 573 TATCGCGGACCGAGTGTCTCTGCTCCGCGCTCAATAGGATATATCCGCGCATAG 514
 Oy 5883 CAGAACTTTAAAGTGTCTCATCATTTGAAAAGCTTCTTCCGCGGCAAACTCTCAAGAT 5942
 Db 513 CAGAACTTTAAAGTGTCTCATCATTTGAAAAGCTTCTTCCGCGGCAAACTCTCAAGAT 454
 Oy 5943 CTACCGCTGTGAGATCCAGTGTGATTAACCCAGTGTGACCCCACTGATCTTCCAGC 6002
 Db 453 CTACCGCTGTGAGATCCAGTGTGATTAACCCAGTGTGACCCCACTGATCTTCCAGC 394
 Oy 6003 ATCTTTACTTTTCAACGAGCTTTCTGCTGAGCAAAAAAGGCAAAATGCGCGCAA 6062
 Db 393 ATCTTTACTTTTCAACGAGCTTTCTGCTGAGCAAAAAAGGCAAAATGCGCGCAA 334
 Oy 6063 AAAGGATATAGGCGGACGACGGAATGTTGAATCTCATCTTCTTCTTCAATATTA 6122
 Db 333 AAAGGATATAGGCGGACGACGGAATGTTGAATCTCATCTTCTTCTTCAATATTA 274
 Oy 6123 TTGAAGCATTTATCAGGTTATTTGCTCATGAGCGGATACATATTTGAATGATTTAGAA 6182
 Db 273 TTGAAGCATTTATCAGGTTATTTGCTCATGAGCGGATACATATTTGAATGATTTAGAA 214
 Oy 6183 AAATTAACAAATAGGCGGTTCCGCGACATTTTCCCGAAAGTGCAC 6229
 Db 213 AAATTAACAAATAGGCGGTTCCGCGACATTTTCCCGAAAGTGCAC 167

RESULT 15
 CD458333/c 870 bp mRNA linear EST 03-JUN-2003
 LOCUS Fg08_09e07_R Fg08_AAFc_ECORC_Fusarium_graminearum_complex_substrate
 DEFINITION Gibberella zeae cDNA clone Fg08_09e07, mRNA sequence.
 ACCESSION CD458333
 VERSION CD458333.1 GI:31373073
 KEYWORDS EST.
 SOURCE
 ORGANISM Gibberella zeae
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
 REFERENCE 1 (bases 1 to 870)
 AUTHORS Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J., Hattori
 J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A., Spott,D. and
 Tinker,N.A.
 TITLE A cDNA library prepared from *Fusarium graminearum* grown on a
 complex plant substrate

JOURNAL Unpublished
 COMMENT Contact: Watson, Robert.J.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-food Canada
 Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
 CANADA
 Tel: (613) 759-1655
 Fax: (613) 759-1701
 Email: watsonrj@agr.gc.ca.

FEATURES
 source
 1..870
 /organism="Gibberella zeae"
 /mol_type="mRNA"
 /strain="DAOM 180378"
 /db_xref="taxon:5518"
 /clone="Fg08_09e07"
 /tissue_type="mycelium"
 /dev_stage="Asexual"
 /lab_host="E. coli DH10B"
 /clone_id="Fg08_AAFc_ECORC_Fusarium_graminearum_complex_s
 ubstrate"

BASE COUNT 239 a 195 c 203 g 230 t 3 others

ORIGIN

Query Match 13.1%; Score 816.8; DB 14; Length 870;
 Best Local Similarity 99.4%; Pred. No. 1.5e-15;
 Matches 837; Conservative 3; Mismatches 0; Indels 2; Gaps 2;

Oy 5389 CGCTACCGGCTCCAGATTTTTCAGCAATTAACACGAGCGGAAAGGCGGAGCGCGA 5448
 Db 870 CGCTACCGGCTCCAGATTTTTCAGCAATTAACACGAGCGGAAAGGCGGAGCGCGA 811
 Oy 5449 AGTGTCTCTGCAACTTATTCGCCCTCATCCAGTCTATTAATGTTGCGGGAAGCTAGA 5508
 Db 810 AGTGTCTCTGCAACTTATTCGCCCTCATCCAGTCTATTAATGTTGCGGGAAGCTAGA 751
 Oy 5509 GTAAGTGTGCGGCAAGTATAGTTTGGGCAACGTTGTCATGTTGCTACAGGATCGTG 5568
 Db 750 GTAAGTGTGCGGCAAGTATAGTTTGGGCAACGTTGTCATGTTGCTACAGGATCGTG 691
 Oy 5569 GTGTACGCTGTGCTGTTGATGATGCTTATTCAGCTCCGCTTCCCAAGATCAAGCGGA 5628
 Db 690 GTGTACGCTGTGCTGTTGATGATGCTTATTCAGCTCCGCTTCCCAAGATCAAGCGGA 631
 Oy 5629 GTTACATGATCCCGCATGTTGTGCAAAAAGCGGTTAGCTCTTCCGCTCGCATGCTT 5688
 Db 630 GTTACATGATCCCGCATGTTGTGCAAAAAGCGGTTAGCT-CTTCCGCTCTCCGATGCTT 572
 Oy 5689 GTACAG-AAGTAAAGTTGGCGGAGTGTATTCATCTCATGTTTGTGCGAGCAGCAATATTC 5747
 Db 571 GTACAGAAAGTAAAGTTGGCGGAGTGTATTCATCTCATGTTTGTGCGAGCAGCAATATTC 512
 Oy 5748 TCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGTGATGATCAACCAAGTC 5807
 Db 511 TCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGTGATGATCAACCAAGTC 452
 Oy 5808 ATTCTGAGATATGATGATGCGGCGACGAGTGTCTTCCGCGGCTCAATAGGATTA 5867
 Db 451 ATTCTGAGATATGATGATGCGGCGACGAGTGTCTTCCGCGGCTCAATAGGATTA 392
 Oy 5868 TACCGGCGCATATGAGCAACTTAAAGTCTCATATCTTGAAGAGCTTTTGGGGCG 5927
 Db 391 TACCGGCGCATATGAGCAACTTAAAGTCTCATATCTTGAAGAGCTTTTGGGGCG 332
 Oy 5928 AAAACTGCAAGGATCTTACCGCTGTGATGATGATGATGATGATGATGATGATGATGATG 5987
 Db 331 AAAACTGCAAGGATCTTACCGCTGTGATGATGATGATGATGATGATGATGATGATGATG 272
 Oy 5988 CAACGTATCTTACAGATCTTTTACTTTCACCGAGGCTTCTGAGAGCAAAAACAGGAG 6047

```

Db      271  CAACTGATCTTCAGCATCTTTACTTTCACCAGCGTTCCTGSGTGAGCAAAAACAGAGAG 212
Qy      6048 GCAAAATGCCGCAAAAAGGGAATAGGGCCGACAGCAAAATGTGAATACTCTACTCTT 6107
Db      211  GCAAAATGCCGCAAAAAGGGAATAGGGCCGACAGCAAAATGTGAATACTCTACTCTT 152
Qy      6108 CCTTTCAATATATATGAGCATTTATCAGGGTTATGTCTCATGAGCGGATACATAT 6167
Db      151  CCTTTCAATATATATGAGCATTTATCAGGGTTATGTCTCATGAGCGGATACATAT 92
Qy      6168 TGAATGTATTTAGAAAAATTAACAATAAGGGGTCGCGCACATTTCCCGAAAAAGTGCC 6227
Db      91  TGAATGTATTTAGAAAAATTAACAATAAGGGGTCGCGCACATTTCCCGAAAAAGTGCC 32
Qy      6228 AC 6229
Db      31  AC 30

```

Search completed: September 17, 2003, 17:25:02
 Job time : 1165 secs

THIS PAGE BLANK (USPTO)